GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 5, 2003, 16:56:07 ; Search time 44 Seconds
(without alignments)
292.773 Million cell updates/sec

Title: US-09-142-524D-3
Perfect score: 711
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....KFIRRVDGIIAAYQNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	188	26.4	374	2	JC2124	major allergen Cry
2	188	26.4	374	2	JC2123	major allergen Cry
3	148.5	20.9	514	2	S48730	Cry j II protein -
4	148.5	20.9	514	2	JC2498	second major aller
5	118.5	16.7	514	2	JC7100	polygalacturonase
6	110.5	15.5	542	2	T06728	pectate lyase (EC
7	107.5	15.1	507	2	JC7366	Jun a 2 protein -
8	104.5	14.7	463	2	T46165	pectate lyase-like
9	103	14.5	404	2	S12209	pectate lyase (EC
10	99	13.9	472	2	T51456	pectate lyase-like
11	98.5	13.9	396	2	A39099	allergen Amb a I.1
12	97	13.6	397	2	C53240	allergen Amb a I.3
13	97	13.6	397	2	C39099	allergen Amb a I.3
14	97	13.6	455	2	T00856	pectate lyase (EC
15	95.5	13.4	459	2	G86278	hypothetical prote
16	94.5	13.3	404	2	T05556	pectate lyase (EC
17	94.5	13.3	431	2	F86179	hypothetical prote
18	93	13.1	397	2	S26211	pectate lyase (EC
19	93	13.1	398	2	B39099	allergen Amb a I.2
20	93	13.1	450	2	T09524	probable pectate l
21	92	12.9	397	2	E53240	allergen Amb a II
22	92	12.9	398	2	B53240	allergen Amb a I.2
23	91.5	12.9	368	2	G86427	probable pectate l
24	91	12.8	398	2	T07058	pectate lyase (EC
25	89	12.5	392	2	D53240	allergen Amb a I.4
26	89	12.5	434	2	S29612	pectate lyase (EC
27	88	12.4	438	2	S43335	pectate lyase (EC
28	84	11.8	374	2	T05240	pectate lyase (EC
29	84	11.8	374	2	H85148	probable pectate l
30	84	11.8	390	2	H86253	hypothetical prote
31	84	11.8	394	2	T49115	pectate lyase like
32	82.5	11.6	449	2	S27098	pectate lyase (EC
33	82	11.5	394	2	T49116	pectate lyase like
34	81	11.4	219	2	C84647	probable synaptobr
35	79	11.1	274	2	T03932	pectate lyase (EC
36	78.5	11.0	418	2	T07701	pectate lyase (EC
37	78	11.0	341	2	T47653	pectate lyase-like
38	76.5	10.8	648	2	S10869	enterotoxin A - Cl
39	76	10.7	398	2	F72335	hypothetical prote
40	74.5	10.5	368	2	F83785	required for glyco
41	74	10.4	219	2	T04630	synaptobrevin homo
42	71.5	10.1	657	2	T22451	hypothetical prote
43	71	10.0	2710	2	A37052	toxin A - Clostrid
44	70.5	9.9	622	2	S17402	parasporal crystal
45	70.5	9.9	1003	2	T26746	hypothetical prote

ALIGNMENTS

RESULT 1
JC2124
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
C;Species: Cryptomeria japonica (Japanese cedar)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C;Accession: JC2124
R;Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A;Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A;Reference number: JC2123; MUID:94183234; PMID:8135802
A;Accession: JC2124
A;Molecule type: mRNA
A;Residues: 1-374 <SON>
A;Cross-references: GB:D26545; NID:g493633; PIDN:BAA05543.1; PID:g493634
A;Experimental source: pollen
A;Note: the authors described carbohydrate binding site for residue 279
C;Superfamily: pectate lyase LAT59
C;Keywords: glycoprotein; pollen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>
F;158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 188; DB 2; Length 374;
Best Local Similarity 71.2%; Pred. No. 1.9e-12;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           ||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138
```

RESULT 2
JC2123
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar
C;Species: Cryptomeria japonica (Japanese cedar)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C;Accession: JC2123; PC2065
R;Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A;Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A;Reference number: JC2123; MUID:94183234; PMID:8135802
A;Accession: JC2123
A;Molecule type: mRNA
A;Residues: 1-374 <SON>
A;Cross-references: GB:D26544; NID:g493631; PIDN:BAA05542.1; PID:g493632
A;Experimental source: pollen
A;Accession: PC2065
A;Molecule type: protein
A;Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO2>
A;Note: the authors described carbohydrate binding site for residue 279
C;Superfamily: pectate lyase LAT59
C;Keywords: glycoprotein; pollen

```
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>
F;158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             26.4%;  Score 188;  DB 2;  Length 374;
  Best Local Similarity   71.2%;  Pred. No. 1.9e-12;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138


RESULT 3
S48730
Cry j II protein - Japanese cedar
C;Species: Cryptomeria japonica (Japanese cedar)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C;Accession: S48730
R;Namba, M.; Kurose, M.; Torigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur
FEBS Lett. 353, 124-128, 1994
A;Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p
A;Reference number: S48730; MUID:95010777; PMID:7926035
A;Accession: S48730
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-514 <NAM>
A;Cross-references: GB:D37765; NID:g577695; PIDN:BAA07021.1; PID:d1007598; PID:g577696

  Query Match             20.9%;  Score 148.5;  DB 2;  Length 514;
  Best Local Similarity   41.4%;  Pred. No. 4.9e-08;
  Matches   41;  Conservative    8;  Mismatches   21;  Indels   29;  Gaps    2;

Qy      31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
           ||||||||||||||||||                          | :|: || ::
Db     236 IDIFASKNFHLQKNTIGTG----------------------DDCVAIGTGSSNIVIEDL 272

Qy      91 MYIAGYKTFDG------RRAEVSYVHVNGAKFIRRVDGI 123
           :   |:    |       ||||||||||||||   :|:
Db     273 ICGPGHGISIGSLGRENSRAEVSYVHVNGAKFIDTQNGL 311


RESULT 4
JC2498
second major allergen Cry j II precursor - Japanese cedar
C;Species: Cryptomeria japonica (Japanese cedar)
C;Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C;Accession: JC2498; PC2346; A60147
R;Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A;Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A;Reference number: JC2498; MUID:94271186; PMID:8002972
A;Accession: JC2498
A;Molecule type: mRNA
A;Residues: 1-514 <KOM>
A;Cross-references: DDBJ:D29772; NID:g506857; PIDN:BAA06172.1; PID:g506858
A;Accession: PC2346
A;Molecule type: protein
A;Residues: 52-61 <KO2>
R;Sakaguchi, M.; Inouye, S.; Taniai, M.; Ando, S.; Usui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A;Title: Identification of the second major allergen of Japanese cedar pollen.
A;Reference number: A60147; MUID:90342988; PMID:2382797
A;Accession: A60147
A;Molecule type: protein
A;Residues: 55-64 <SAK>
C;Keywords: glycoprotein; pollen
F;1-54/Domain: signal sequence #status predicted <SIG>
F;55-460/Product: second major allergen Cry j #status predicted <MAT>
F;429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             20.9%;  Score 148.5;  DB 2;  Length 514;
  Best Local Similarity   41.4%;  Pred. No. 4.9e-08;
  Matches   41;  Conservative    8;  Mismatches   21;  Indels   29;  Gaps    2;

Qy      31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
           ||||||||||||||||||                          | :|: || ::
Db     236 IDIFASKNFHLQKNTIGTG----------------------DDCVAIGTGSSNIVIEDL 272

Qy      91 MYIAGYKTFDG------RRAEVSYVHVNGAKFIRRVDGI 123
           :   |:    |       ||||||||||||||   :|:
Db     273 ICGPGHGISIGSLGRENSRAEVSYVHVNGAKFIDTQNGL 311


RESULT 5
JC7100
polygalacturonase Cha o 2 - Japanese cypress
C;Species: Chamaecypris obtusa (Japanese cypress)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7100; PC7026
R;Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A;Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A;Reference number: JC7100; MUID:99417540; PMID:10486272
A;Accession: JC7100
A;Molecule type: mRNA
A;Residues: 1-514 <MOR>
A;Accession: PC7026
A;Molecule type: protein
A;Residues: 51-62 <MO2>

  Query Match             16.7%;  Score 118.5;  DB 2;  Length 514;
  Best Local Similarity   35.4%;  Pred. No. 8.3e-05;
  Matches   35;  Conservative   10;  Mismatches   25;  Indels   29;  Gaps    2;

Qy      31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
           |||||||  ||::|  ||||                          | :|: || :|
Db     236 IDIFASKRFHIEKCVIGTG----------------------DDCIAIGTGSSNITIKDL 272

Qy      91 MYIAGYKTF------DGRRAEVSYVHVNGAKFIRRVDGI 123
           :   |:         |  |||||:|||| ||||   :|:
Db     273 ICGPGHGISIGSLGRDNSRAEVSHVHVNRAKFIDTQNGL 311


RESULT 6
T06728
pectate lyase (EC 4.2.2.2) F28P10.100 - Arabidopsis thaliana
N;Alternate names: protein F28P10.100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C;Accession: T06728
R;Quetier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T06728
A;Molecule type: DNA
A;Residues: 1-542 <QUE>
A;Cross-references: EMBL:AL049655
A;Experimental source: cultivar Columbia; BAC clone F28P10
C;Genetics:
A;Map position: 3
A;Introns: 46/2; 346/3; 413/2; 480/2
A;Note: F28P10.100
C;Superfamily: pectate lyase LAT59
C;Keywords: carbon-oxygen lyase

  Query Match             15.5%;  Score 110.5;  DB 2;  Length 542;
  Best Local Similarity   43.3%;  Pred. No. 0.00064;
  Matches   26;  Conservative    9;  Mismatches   20;  Indels    5;  Gaps    2;

Qy      67 FTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV--NGAKFIRRVDGII 124
           : ::   ||||:|| || |:||  : |  ||| |||    | ||:  ||   :: |  ||
Db     139 YGVIQEEPLWIVFSSNMLIRLKQELIINSYKTLDGRG---SAVHITGNGCLTLQYVQHII 195


RESULT 7
```

JC7366
Jun a 2 protein - mountain cedar
C;Species: Juniperus ashei (mountain cedar)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C;Accession: JC7366; PC7093
R;Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.
Biochem. Biophys. Res. Commun. 275, 195-202, 2000
A;Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all
A;Reference number: JC7366
A;Accession: JC7366
A;Molecule type: mRNA
A;Residues: 1-507 <YOK>
A;Cross-references: GB:AJ404653
A;Accession: PC7093
A;Molecule type: protein
A;Residues: 55-63 <YO2>
C;Comment: This protein, a second major allergen of mountain cedar pollen, which is invo
o the polygalacturonase family.
C;Keywords: glycoprotein; pollen

Query Match 15.1%; Score 107.5; DB 2; Length 507;
Best Local Similarity 32.3%; Pred. No. 0.0012;
Matches 32; Conservative 13; Mismatches 25; Indels 29; Gaps 2;

Qy 31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
Db 237 IDIFASKRFEIEKCTIGTG----------DDCVAVGTGSSNITIKDL 273

Qy 91 MYIAGYKTFDG------RRAEVSYVHVNGAKFIRRVDGI 123
Db 274 TCGPGHGMSIGSLGKGNSRSEVSFVHLDGAKFIDTQNGL 312

RESULT 8
T46165
pectate lyase-like protein - Arabidopsis thaliana
N;Alternate names: protein T4D2.120
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C;Accession: T46165
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23025
A;Accession: T46165
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-463 <NYA>
A;Cross-references: EMBL:AL132958
A;Experimental source: cultivar Columbia; BAC clone T4D2
C;Genetics:
A;Map position: 3
A;Introns: 27/2; 112/2; 182/1; 265/3; 329/3; 399/2
A;Note: T4D2.120
C;Superfamily: pectate lyase LAT59

Query Match 14.7%; Score 104.5; DB 2; Length 463;
Best Local Similarity 30.5%; Pred. No. 0.0024;
Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;

Qy 52 WKNNRIWL----------QFAKLTG----------FTLMGRRPLWIIF 79
Db 66 WPNNRQGLADCGIGFGQYALGGKGGQFYFVTDSSDDDAVNPKPGTLRYGVIQEEPLWIVF 125

Qy 80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
Db 126 PSNMMIKLKQELIFNSYKTLDGRGANVHIVG-GGCITLQYVSNII 169

RESULT 9
S12209
pectate lyase (EC 4.2.2.2) - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C;Accession: S12209
R;Budelier, K.A.; Smith, A.G.; Gasser, C.S.
Mol. Gen. Genet. 224, 183-192, 1990
A;Title: Regulation of a stylar transmitting tissue-specific gene in wild-type and trans
A;Reference number: S12209; MUID:91117185; PMID:2277637
A;Accession: S12209
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-404 <BUD>
A;Cross-references: GB:X55193; NID:g19161; PIDN:CAA38979.1; PID:g19162
C;Superfamily: pectate lyase LAT59
C;Keywords: carbon-oxygen lyase

Query Match 14.5%; Score 103; DB 2; Length 404;
Best Local Similarity 30.8%; Pred. No. 0.0029;
Matches 33; Conservative 13; Mismatches 31; Indels 30; Gaps 4;

Qy 28 GRRIDIFASKNFHLQKNTIGTGRRWKNNRIW----------LQFAKLTGFTLM 70
Db 67 GKKSPAFSYCAIGFGKNAIG----GKNGRIYVVTDSGNDDPVNPKPGTLRHA------VI 116

Qy 71 GRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFI 117
Db 117 QDEPLWIIFKRDMVIQLKQELVMNSYKTIDGRGAS---VHISGGPCI 160

RESULT 10
T51456
pectate lyase-like protein - Arabidopsis thaliana
N;Alternate names: protein F2G14_230
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
C;Accession: T51456
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51456
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <SAT>
A;Cross-references: EMBL:AL391146
A;Experimental source: cultivar Columbia; BAC clone F2G14
C;Genetics:
A;Map position: 5
A;Introns: 65/2; 316/3; 398/3
A;Note: F2G14_230
C;Superfamily: pectate lyase LAT59

Query Match 13.9%; Score 99; DB 2; Length 472;
Best Local Similarity 28.2%; Pred. No. 0.0094;
Matches 29; Conservative 17; Mismatches 41; Indels 16; Gaps 3;

Qy 34 FASKNFHLQKNTIGTGRR---WKNNRIWLQFAKL----------TGFTLMGRRPLWIIF 79
Db 134 WAKRRKKLTRCVRGFGHRTTGGKRGRIYVVTSNLDEDMVNPKPGTLRHAVIQKEPLWIIF 193

Qy 80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNG--AKFIRRV 120
Db 194 KNDMSIRLNQELLINSHKTIDARGANVHVAHGAGITMQFVKNV 236

RESULT 11
A39099
allergen Amb a I.1 precursor - common ragweed
C;Species: Ambrosia artemisiifolia (common ragweed)
C;Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
C;Accession: A39099; A60895; A53240
R;Rafnar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
J. Biol. Chem. 266, 1229-1236, 1991
A;Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed poll
A;Reference number: A39099; MUID:91093235; PMID:1702434
A;Accession: A39099
A;Molecule type: mRNA

```
A;Residues: 1-396 <RAF>
A;Cross-references: GB:M63116
R;Smith, J.J.; Olson, J.R.; Klapper, D.G.
Mol. Immunol. 25, 355-365, 1988
A;Title: Monoclonal antibodies to denatured ragweed pollen allergen Amb a I: characteriz
 a I.
A;Reference number: A60895; MUID:88288254; PMID:2456454
A;Accession: A60895
A;Molecule type: protein
A;Residues: 256-273;292-303,'W',305-306 <SMI>
R;Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
A;Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
A;Reference number: A53240; MUID:92234570; PMID:1809687
A;Accession: A53240
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-396 <GRI>
A;Cross-references: GB:M80558; NID:g166434; PIDN:AAA32665.1; PID:g166435
C;Superfamily: pectate lyase LAT59
C;Keywords: glycoprotein; pollen
F;36/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             13.9%;  Score 98.5;  DB 2;  Length 396;
  Best Local Similarity   34.7%;  Pred. No. 0.0087;
  Matches   26;  Conservative   12;  Mismatches   22;  Indels   15;  Gaps    3;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVH----VNGAK--FIRRVD----- 121
           |||||||  :| |:|   | :   || ||| |:|  ::    :|| |   |  ::
Db     113 RPLWIIFERDMVIRLDKEMVVNSDKTIDGRGAKVEIINAGFTLNGVKNVIIHNINMHDVK 172

Qy     122 ----GIIAAYQNPAS 132
               |:| :   ||:
Db     173 VNPGGLIKSNDGPAA 187


RESULT 12
C53240
allergen Amb a I.3 precursor - common ragweed
C;Species: Ambrosia artemisiifolia (common ragweed)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 18-Sep-1998
C;Accession: C53240
R;Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
A;Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
A;Reference number: A53240; MUID:92234570; PMID:1809687
A;Accession: C53240
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-397 <GRI>
A;Cross-references: GB:M80560
C;Superfamily: pectate lyase LAT59
C;Keywords: pollen

  Query Match             13.6%;  Score 97;  DB 2;  Length 397;
  Best Local Similarity   29.5%;  Pred. No. 0.013;
  Matches   31;  Conservative    7;  Mismatches   33;  Indels   34;  Gaps    3;

Qy      52 WKNNRIWLQ-----FAKLT-------------------------GFTLMGRRPLWIIF 79
           |:|||  |      ||| |                         |     |||||||
Db      61 WENNRQALADCAQGFAKGTYGGKWGDVYTVTSNLDDDVANPKEGTLRFAAAQNRPLWIIF 120

Qy      80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
             :| | |   : :   || |||  :|    :||   :  |  ||
Db     121 KNDMVINLNQELVVNSDKTIDGRGVKVEI--INGGLTLMNVKNII 163


RESULT 13
C39099
allergen Amb a I.3 - common ragweed
C;Species: Ambrosia artemisiifolia (common ragweed)
C;Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
C;Accession: C39099
R;Rafnar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
J. Biol. Chem. 266, 1229-1236, 1991
A;Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed poll
A;Reference number: A39099; MUID:91093235; PMID:1702434
A;Accession: C39099
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-397 <RAF>
A;Cross-references: GB:M62961; NID:g166440; PIDN:AAA32668.1; PID:g166441
C;Superfamily: pectate lyase LAT59
C;Keywords: pollen

  Query Match             13.6%;  Score 97;  DB 2;  Length 397;
  Best Local Similarity   29.5%;  Pred. No. 0.013;
  Matches   31;  Conservative    7;  Mismatches   33;  Indels   34;  Gaps    3;

Qy      52 WKNNRIWLQ-----FAKLT-------------------------GFTLMGRRPLWIIF 79
           |:|||  |      ||| |                         |     |||||||
Db      61 WENNRQALADCAQGFAKGTYGGKWGDVYTVTSNLDDDVANPKEGTLRFAAAQNRPLWIIF 120

Qy      80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
             :| | |   : :   || |||  :|    :||   :  |  ||
Db     121 KNDMVINLNQELVVNSDKTIDGRGVKVEI--INGGLTLMNVKNII 163


RESULT 14
T00856
pectate lyase (EC 4.2.2.2) T20F6.14 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C;Accession: T00856; A84440
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, March 1998
A;Description: Arabidopsis thaliana chromosome II BAC T20F6 genomic sequence.
A;Reference number: Z14206
A;Accession: T00856
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-455 <ROU>
A;Cross-references: EMBL:AC002521; NID:g2947056; PIDN:AAC05350.1; PID:g2947069
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-455 <STO>
A;Cross-references: GB:AE002093; NID:g2947069; PIDN:AAC05350.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02720
A;Map position: 2
A;Introns: 66/2; 295/3; 376/3
A;Note: T20F6.14
C;Superfamily: pectate lyase LAT59
C;Keywords: carbon-oxygen lyase

  Query Match             13.6%;  Score 97;  DB 2;  Length 455;
  Best Local Similarity   46.2%;  Pred. No. 0.015;
  Matches   24;  Conservative    9;  Mismatches   13;  Indels    6;  Gaps    3;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGA----KFIRRV 120
           |||||||: :| |||:  : |    || ||| |:: |:   ||    :|:| |
Db     166 RPLWIIFARSMIIKLQQELIITNDKTIDGRGAKI-YI-TGGAGLTLQFVRNV 215


RESULT 15
G86278
hypothetical protein F14L17.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
```

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 01-Mar-2002
C;Accession: G86278
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86278
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <STO>
A;Cross-references: GB:AE005172; NID:g7262684; PIDN:AAF43942.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: pectate lyase LAT59

Query Match 13.4%; Score 95.5; DB 2; Length 459;
Best Local Similarity 28.6%; Pred. No. 0.022;
Matches 32; Conservative 9; Mismatches 28; Indels 43; Gaps 4;

```
Qy      50 RRWKNNRIWLQFAKLTGFTL-MGRR---------------------------------- 73
           | | |||      ||    |  |||
Db     115 RNWANNR-----KKLADCVLGFGRRTTGGKDGPIYVVKDASDNDLINPKPGTLRHAVTRD 169

Qy      74 -PLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
            ||||||: :| |||:  : |   || ||| | | |:       :: |: :|
Db     170 GPLWIIFARSMIIKLQQELMITSDKTIDGRGARV-YIMEGAGLTLQFVNNVI 220
```

Search completed: May 5, 2003, 16:59:58
Job time : 55 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 5, 2003, 16:48:41 ; Search time 12 Seconds
(without alignments)
463.152 Million cell updates/sec

Title: US-09-142-524D-3
Perfect score: 711
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....KFIRRVDGIIAAYQNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	188	26.4	374	1	SBP_CRYJA	P18632 cryptomeria
2	150	21.1	367	1	MPA1_JUNAS	P81294 juniperus a
3	148.5	20.9	514	1	MPA2_CRYJA	P43212 cryptomeria
4	143	20.1	346	1	MPA1_CUPAR	Q9scg9 cupressus a
5	142	20.0	375	1	MPA1_CHAOB	Q96385 chamaecypar
6	103	14.5	404	1	9612_LYCES	P24396 lycopersico
7	98.5	13.9	396	1	MP11_AMBAR	P27759 ambrosia ar
8	97	13.6	397	1	MP13_AMBAR	P27761 ambrosia ar
9	93	13.1	397	1	PEL_TOBAC	P40972 nicotiana t
10	93	13.1	398	1	MP12_AMBAR	P27760 ambrosia ar
11	92	12.9	397	1	MPA2_AMBAR	P27762 ambrosia ar
12	91	12.8	398	1	PE56_LYCES	P15721 lycopersico
13	89	12.5	392	1	MP14_AMBAR	P28744 ambrosia ar
14	89	12.5	434	1	PEL_LILLO	P40973 lilium long
15	82.5	11.6	449	1	PE59_LYCES	P15722 lycopersico
16	81.5	11.5	234	1	A29B_DROME	O46197 drosophila
17	80.5	11.3	605	1	PHAC_METEX	P52070 methylobact
18	71	10.0	2710	1	TOXA_CLODI	P16154 clostridium
19	70.5	9.9	622	1	C2AC_BACTU	Q45743 bacillus th
20	69.5	9.8	338	1	YJX8_YEAST	P47085 saccharomyc
21	68.5	9.6	234	1	A29B_DROSI	Q9u968 drosophila
22	68.5	9.6	551	1	CHIT_NPVAC	P41684 autographa
23	67.5	9.5	985	1	AGLU_ASPOR	Q12558 aspergillus
24	67.5	9.5	1070	1	YHV4_YEAST	P38850 saccharomyc
25	67	9.4	757	1	FIXI_RHIME	P18398 rhizobium m
26	66.5	9.4	110	1	Y12K_SMSV4	P36289 san miguel
27	66	9.3	375	1	Y4ED_RHISN	P55427 rhizobium s
28	66	9.3	382	1	PHYT_BACSU	P42094 bacillus su
29	66	9.3	646	1	NODQ_RHIS3	P72339 r nodq bifu
30	65	9.1	576	1	DPEP_SOLTU	Q06801 solanum tub
31	65	9.1	726	1	YB1D_SCHPO	P87178 schizosacch
32	65	9.1	781	1	TL22_CHICK	Q9dgb6 gallus gall
33	64.5	9.1	1010	1	CLPP_CHLEU	P42379 chlamydomon
34	64	9.0	532	1	PUR2_ARATH	P52420 arabidopsis
35	64	9.0	633	1	NODQ_RHISB	O07309 r nodq bifu
36	63.5	8.9	196	1	YM07_YEAST	Q04487 saccharomyc
37	63.5	8.9	309	1	Y4OQ_RHISN	P55602 rhizobium s
38	63.5	8.9	326	1	SCRR_PEDPE	P43472 pediococcus
39	63	8.9	347	1	NU2M_RABIT	O79428 oryctolagus
40	63	8.9	477	1	GLGA_SYNY3	P74521 synechocyst
41	63	8.9	522	1	CPV1_ORENI	P70091 oreochromis
42	63	8.9	725	1	GUNG_CLOCE	P37700 clostridium
43	63	8.9	755	1	COMP_RAT	P35444 rattus norv
44	63	8.9	1121	1	WDR6_HUMAN	Q9nnw5 homo sapien
45	62.5	8.8	262	1	FLGG_AGRT5	Q44338 agrobacteri

ALIGNMENTS

RESULT 1
SBP_CRYJA
ID SBP_CRYJA STANDARD; PRT; 374 AA.
AC P18632;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j I).
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=94183234; PubMed=8135802;
RA Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,
RA Kino K.;
RT "Cloning and sequencing of cDNA coding for Cry j I, a major allergen
RT of Japanese cedar pollen.";
RL Biochem. Biophys. Res. Commun. 199:619-625(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RA Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 22-41.
RC TISSUE=Pollen;
RX MEDLINE=89031257; PubMed=3181436;
RA Taniai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S.,
RA Matuhasi T.;
RT "N-terminal amino acid sequence of a major allergen of Japanese cedar
RT pollen (Cry j I).";
RL FEBS Lett. 239:329-332(1988).
RN [4]
RP CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=95003748; PubMed=7920021;
RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.;
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
RT (Cryptomeria japonica) pollen allergen, Cry jI.";
RL Int. Arch. Allergy Immunol. 105:198-202(1994).
RN [5]
RP STRUCTURE OF CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=95332249; PubMed=7608114;
RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,
RA Fukuda S., Hanzawa H., Haruyama H., Kurimoto M.;
RT "Carbohydrate structures of the glycoprotein allergen Cry j I from
RT Japanese cedar (Cryptomeria japonica) pollen.";
RL J. Biochem. 117:289-295(1995).
CC -!- PTM: CONTAINS FUCOSE/XYLOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC -!- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
CC POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM

```
CC       B DIFFERS IN SIX POSITIONS.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D26544; BAA05542.1; -.
DR   EMBL; D26545; BAA05543.1; -.
DR   EMBL; D34639; BAA07020.1; -.
DR   PIR; A44773; A44773.
DR   GlycoSuiteDB; P18632; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Allergen; Glycoprotein; Multigene family; Signal.
FT   SIGNAL       1     21
FT   CHAIN       22    374       SUGI BASIC PROTEIN.
FT   CARBOHYD   158    158       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   191    191       N-LINKED (GLCNAC. . .) (COMPLEX).
FT                               /FTId=CAR_000135.
FT   CARBOHYD   293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   354    354       N-LINKED (GLCNAC. . .) (COMPLEX).
FT                               /FTId=CAR_000136.
FT   VARIANT     12     12       L -> F (IN CRY J 1-B).
FT   VARIANT    143    143       H -> Y (IN CRY J 1-B).
FT   VARIANT    202    202       S -> T (IN CRY J 1-B).
FT   VARIANT    221    221       L -> S (IN CRY J 1-B).
FT   VARIANT    358    358       Q -> H (IN CRY J 1-B).
FT   VARIANT    361    361       K -> Q (IN CRY J 1-B).
SQ   SEQUENCE   374 AA;  40645 MW;  74AB25950248F56F CRC64;

  Query Match             26.4%;  Score 188;  DB 1;  Length 374;
  Best Local Similarity   71.2%;  Pred. No. 2.7e-12;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||||| |:|    :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138


RESULT 2
MPA1_JUNAS
ID   MPA1_JUNAS     STANDARD;      PRT;   367 AA.
AC   P81294; Q9ZNU7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Major pollen allergen Jun a 1 precursor.
OS   Juniperus ashei (Ozark white cedar).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX   NCBI_TaxID=13101;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
RP   117-124; 134-140; 160-164; 256-263 AND 322-325.
RC   TISSUE=Pollen;
RX   MEDLINE=99414163; PubMed=10482836;
RA   Midoro-Horiuti T.M., Goldblum R.M., Kurosky A., Wood T.G.,
RA   Brooks E.G.;
RT   "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
RT   allergen, Jun a 1.";
RL   J. Allergy Clin. Immunol. 104:613-617(1999).
RN   [2]
RP   SEQUENCE OF 22-50.
RC   TISSUE=Pollen;
RX   MEDLINE=99414162; PubMed=10482835;
RA   Midoro-Horiuti T., Goldblum R.M., Kurosky A., Goetz D.W.,
RA   Brooks E.G.;
RT   "Isolation and characterization of the mountain cedar (Juniperus
RT   ashei) pollen major allergen, Jun a 1.";
RL   J. Allergy Clin. Immunol. 104:608-612(1999).
CC   -!- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIS RHINITIS
CC       IN NORTH AMERICA.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF106663; AAD03609.1; -.
DR   EMBL; AF106662; AAD03608.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Allergen; Glycoprotein; Signal.
FT   SIGNAL       1     21
FT   CHAIN       22    367       MAJOR POLLEN ALLERGEN JUN A 1.
FT   CARBOHYD   148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   367 AA;  39824 MW;  FC9B81E675662E49 CRC64;

  Query Match             21.1%;  Score 150;  DB 1;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 2.4e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||||:|:||:|| ||| |:|    :     |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 3
MPA2_CRYJA
ID   MPA2_CRYJA     STANDARD;      PRT;   514 AA.
AC   P43212;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE   (Major pollen allergen Cry j 2) (Cry j II).
OS   Cryptomeria japonica (Japanese cedar).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX   NCBI_TaxID=3369;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Pollen;
RX   MEDLINE=95010777; PubMed=7926035;
RA   Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,
RA   Usui M., Kurimoto M.;
RT   "Molecular cloning of the second major allergen, Cry j II, from
RT   Japanese cedar pollen.";
RL   FEBS Lett. 353:124-128(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pollen;
RX   MEDLINE=94271186; PubMed=8002972;
RA   Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;
RT   "cDNA cloning and expression of Cry j II the second major allergen of
RT   Japanese cedar pollen.";
RL   Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN   [3]
RP   SEQUENCE OF 55-64.
RX   MEDLINE=90342988; PubMed=2382797;
RA   Sakaguchi M., Inouye S., Taniai M., Ando S., Usui M., Matuhasi T.;
RT   "Identification of the second major allergen of Japanese cedar
```

```
RT   pollen.";
RL   Allergy 45:309-312(1990).
CC   -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC       galactosiduronic linkages in pectate and other galacturonans.
CC   -!- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC       (POLYGALACTURONASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D37765; BAA07021.1; -.
DR   EMBL; D29772; BAA06172.1; -.
DR   HSSP; P26509; 1BHE.
DR   InterPro; IPR000743; GH28.
DR   Pfam; PF00295; Glyco_hydro_28; 1.
DR   PROSITE; PS00502; POLYGALACTURONASE; 1.
KW   Hydrolase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
KW   Amyloplast; Glycoprotein; Allergen.
FT   SIGNAL       1     22       POTENTIAL.
FT   PROPEP       23    45       POTENTIAL.
FT   CHAIN        46    433      POSSIBLE POLYGALACTURONASE.
FT   PROPEP       434   514      POTENTIAL.
FT   ACT_SITE     278   278      PROBABLE.
FT   CARBOHYD     460   460      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     472   472      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     5     5        F -> L (IN REF. 2).
FT   CONFLICT     12    12       V -> L (IN REF. 2).
FT   CONFLICT     34    35       DI -> VV (IN REF. 2).
FT   CONFLICT     37    37       Q -> K (IN REF. 2).
FT   CONFLICT     88    88       K -> N (IN REF. 2).
FT   CONFLICT     98    98       N -> S (IN REF. 2).
FT   CONFLICT     451   451      K -> E (IN REF. 2).
FT   CONFLICT     454   454      G -> R (IN REF. 2).
FT   CONFLICT     504   504      M -> I (IN REF. 2).
FT   CONFLICT     507   507      R -> C (IN REF. 2).
SQ   SEQUENCE   514 AA;  56645 MW;  624611C3FA8D6302 CRC64;

  Query Match             20.9%;  Score 148.5;  DB 1;  Length 514;
  Best Local Similarity   41.4%;  Pred. No. 4.9e-08;
  Matches   41;  Conservative    8;  Mismatches   21;  Indels   29;  Gaps    2;

Qy      31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
           |||||||||||||||||                                 | :|: || ::
Db     236 IDIFASKNFHLQKNTIGTG----------------------DDCVAIGTGSSNIVIEDL 272

Qy      91 MYIAGYKTFDG------RRAEVSYVHVNGAKFIRRVDGI 123
           :   |:    |       |||||||||||||||   :|:
Db     273 ICGPGHGISIGSLGRENSRAEVSYVHVNGAKFIDTQNGL 311


RESULT 4
MPA1_CUPAR
ID   MPA1_CUPAR     STANDARD;      PRT;   346 AA.
AC   Q9SCG9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Major pollen allergen Cup a 1.
OS   Cupressus arizonica.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=49011;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20571526; PubMed=11122214;
RA   Aceituno E., Del Pozo V., Minguez A., Arrieta I., Cortegano I.,
RA   Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;
RT   "Molecular cloning of major allergen from Cupressus arizonica pollen:
RT   Cup a 1.";
RL   Clin. Exp. Allergy 30:1750-1758(2000).
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ243570; CAB62551.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Allergen; Glycoprotein.
FT   CARBOHYD     127   127      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     157   157      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     272   272      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   346 AA;  37589 MW;  F1281DCDA1D5DFD0 CRC64;

  Query Match             20.1%;  Score 143;  DB 1;  Length 346;
  Best Local Similarity   53.8%;  Pred. No. 1.2e-07;
  Matches   28;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| ||||||:||:|:||||| ||| | |   :     |:|:   :|
Db      66 KALWIIFSQNMNIKLQMPLYVAGYKTIDGRGAVVHLGNGGPCLFMRKASHVI 117


RESULT 5
MPA1_CHAOB
ID   MPA1_CHAOB     STANDARD;      PRT;   375 AA.
AC   Q96385;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Major pollen allergen Cha o 1 precursor.
OS   Chamaecyparis obtusa (Japanese cypress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC   Chamaecyparis.
OX   NCBI_TaxID=13415;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Pollen;
RX   MEDLINE=96265194; PubMed=8676896;
RA   Suzuki M., Komiyama N., Itoh M., Itoh H., Sone T., Kuno K., Takagi I.,
RA   Ohta N.;
RT   "Purification, characterization and molecular cloning of Cha o 1, a
RT   major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
RL   Mol. Immunol. 33:451-460(1996).
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D45404; BAA08246.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Allergen; Glycoprotein; Signal.
FT   SIGNAL       1     21
FT   CHAIN        22    375      MAJOR POLLEN ALLERGEN CHA O 1.
```

```
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   375 AA;  40258 MW;  81CD91DF7066DBBF CRC64;

  Query Match          20.0%;  Score 142;  DB 1;  Length 375;
  Best Local Similarity   57.7%;  Pred. No. 1.7e-07;
  Matches   30;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           | ||||||-|:|||| ||:|||| || ||| |||   :     |:| |  :|
Db      87 RSLWIIFSKNLNIKLNMPLYIAGNKTIDGRGAEVHIGNGGPCLFMRTVSHVI 138


RESULT 6
9612_LYCES
ID   9612_LYCES     STANDARD;      PRT;   404 AA.
AC   P24396;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Style development-specific protein 9612 precursor.
GN   9612.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VF36; TISSUE=Pistil;
RX   MEDLINE=91117185; PubMed=2277637;
RA   Budelier K.A., Smith A.G., Gasser C.S.;
RT   "Regulation of a stylar transmitting tissue-specific gene in
RT   wild-type and transgenic tomato and tobacco.";
RL   Mol. Gen. Genet. 224:183-192(1990).
CC   -!- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
CC       TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
CC       SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
CC       GROWTH.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
CC       FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
CC       TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
CC       LOWER LEVELS IN THE ANTHERS AND VEGETATIVE ORGANS.
CC   -!- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHESIS.
CC   -!- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)
CC       AND P56 (AC P15721).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X55193; CAA38979.1; -.
DR   PIR; S12209; S12209.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    404       STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612.
FT   CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   404 AA;  44298 MW;  B26ED69B128D8675 CRC64;

  Query Match          14.5%;  Score 103;  DB 1;  Length 404;
  Best Local Similarity   30.8%;  Pred. No. 0.0021;
  Matches   33;  Conservative   13;  Mismatches   31;  Indels   30;  Gaps    4;

Qy      28 GRRIDIFASKNFHLQKNTIGTGRRWKNNRIW----------------LQFAKLTGFTLM 70
           |::   |:        || ||     || ||:                |: |      ::
Db      67 GKKSPAFSYCAIGFGKNAIG----GKNGRIYVVTDSGNDDPVNPKPGTLRHA------VI 116

Qy      71 GRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFI 117
           ||||||  :| |:||  : :  ||| ||| |    ||::|   |
Db     117 QDEPLWIIFKRDMVIQLKQELVMNSYKTIDGRGAS---VHISGGPCI 160


RESULT 7
MP11_AMBAR
ID   MP11_AMBAR     STANDARD;      PRT;   396 AA.
AC   P27759;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pollen allergen Amb a 1.1 precursor (Antigen E) (Antigen Amb a I).
OS   Ambrosia artemisiifolia (Short ragweed).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC   Heliantheae; Ambrosia.
OX   NCBI_TaxID=4212;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Pollen;
RX   MEDLINE=91093235; PubMed=1702434;
RA   Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
RA   Klapper D.G.;
RT   "Cloning of Amb a I (antigen E), the major allergen family of short
RT   ragweed pollen.";
RL   J. Biol. Chem. 266:1229-1236(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   TISSUE=Pollen;
RX   MEDLINE=92234570; PubMed=1809687;
RA   Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT   "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT   in Ambrosia artemisiifolia (short ragweed).";
RL   Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M63116; -; NOT_ANNOTATED_CDS.
DR   EMBL; M80558; AAA32665.1; -.
DR   PIR; A39099; A39099.
DR   PIR; A53240; A53240.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    396       POLLEN ALLERGEN AMB A 1.1.
FT   VARIANT      92     92       E -> D.
SQ   SEQUENCE   396 AA;  42709 MW;  0CE7DDECB2B8841D CRC64;

  Query Match          13.9%;  Score 98.5;  DB 1;  Length 396;
  Best Local Similarity   34.7%;  Pred. No. 0.006;
  Matches   26;  Conservative   12;  Mismatches   22;  Indels   15;  Gaps    3;
```

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVH----VNGAK--FIRRVD----- 121
           ||||||| :| |:|   | :   || ||| |:|  ::    :|| |   |  ::
Db     113 RPLWIIFERDMVIRLDKEMVVNSDKTIDGRGAKVEIINAGFTLNGVKNVIIHNINMHDVK 172

Qy     122 ----GIIAAYQNPAS 132
               |:| :   ||:
Db     173 VNPGGLIKSNDGPAA 187
```

RESULT 8
MP13_AMBAR
ID MP13_AMBAR STANDARD; PRT; 397 AA.
AC P27761;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I).
OS Ambrosia artemisiifolia (Short ragweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Ambrosia.
OX NCBI_TaxID=4212;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=91093235; PubMed=1702434;
RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
RA Klapper D.G.;
RT "Cloning of Amb a I (antigen E), the major allergen family of short
RT ragweed pollen.";
RL J. Biol. Chem. 266:1229-1236(1991).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=Pollen;
RX MEDLINE=92234570; PubMed=1809687;
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT in Ambrosia artemisiifolia (short ragweed).";
RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC -!- SUBUNIT: MONOMER.
CC -!- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -!- PTM: The N-terminus is blocked.
CC -!- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC   AMB A I/AMB A II/CRY J I SUBFAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; M62961;-AAA32668.1; -.
DR EMBL; M80560; AAA32669.1; ALT_SEQ.
DR PIR; C39099; C39099.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT SIGNAL   1   25   POTENTIAL.
FT CHAIN   26   397   POLLEN ALLERGEN AMB A 1.3.
FT VARIANT   48   48   L -> Y.
SQ SEQUENCE 397 AA; 42928 MW; C8DB41257590DD0A CRC64;

  Query Match   13.6%; Score 97; DB 1; Length 397;
  Best Local Similarity 29.5%; Pred. No. 0.0086;
  Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;

```
Qy      52 WKNNRIWLQ-----FAKLT-------------------------GFTLMGRRPLWIIF 79
           |:|||  |      ||| |                         |     |||||||
Db      61 WENNRQALADCAQGFAKGTYGGKWGDVYTVTSNLDDDVANPKEGTLRFAAAQNRPLWIIF 120

Qy      80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
             :| | |   : :   || |||  :|   :||   :  |  ||
Db     121 KNDMVINLNQELVVNSDKTIDGRGVKVEI--INGGLTLMNVKNII 163
```

RESULT 9
PEL_TOBAC
ID PEL_TOBAC STANDARD; PRT; 397 AA.
AC P40972;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Pectate lyase precursor (EC 4.2.2.2).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Samsun; TISSUE=Pollen;
RX MEDLINE=93043039; PubMed=1421152;
RA Rogers H.J., Harvey A., Lonsdale D.M.;
RT "Isolation and characterization of a tobacco gene with homology to
RT pectate lyase which is specifically expressed during
RT microsporogenesis.";
RL Plant Mol. Biol. 20:493-502(1992).
CC -!- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
CC   oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
CC   their non-reducing ends.
CC -!- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
CC   DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X67158; CAA47630.1; -.
DR EMBL; X67159; CAA47631.1; -.
DR EMBL; X61102; CAA43414.1; -.
DR PIR; S26211; S26211.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Lyase; Signal.
FT SIGNAL   1   25   POTENTIAL.
FT CHAIN   26   397   PECTATE LYASE.
FT ACT_SITE 272   272   POTENTIAL.
FT CARBOHYD 134   134   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227   227   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 156   156   S -> C (IN MRNA).
FT CONFLICT 189   190   GS -> SG (IN MRNA).
FT CONFLICT 200   200   S -> D (IN MRNA).
FT CONFLICT 202   202   H -> R (IN MRNA).
FT CONFLICT 249   249   H -> N (IN MRNA).
SQ SEQUENCE 397 AA; 44351 MW; EF0A82CE5DA7643F CRC64;

  Query Match   13.1%; Score 93; DB 1; Length 397;
  Best Local Similarity 42.9%; Pred. No. 0.022;
  Matches 21; Conservative 7; Mismatches 17; Indels 4; Gaps 2;

```
Qy      67 FTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV-NGA 114
           : :: : ||||||  || |||   : :   || |||      ||: |||
Db     105 YGVIQKEPLWIIFGKNMKIKLSRELIVTSNKTIDGRGFN---VHIQNGA 150
```

```
RESULT 10
MP12_AMBAR
ID   MP12_AMBAR     STANDARD;      PRT;   398 AA.
AC   P27760;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pollen allergen Amb a 1.2 precursor (Antigen E) (Antigen Amb a I).
OS   Ambrosia artemisiifolia (Short ragweed).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC   Heliantheae; Ambrosia.
OX   NCBI_TaxID=4212;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pollen;
RX   MEDLINE=91093235; PubMed=1702434;
RA   Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
RA   Klapper D.G.;
RT   "Cloning of Amb a I (antigen E), the major allergen family of short
RT   ragweed pollen.";
RL   J. Biol. Chem. 266:1229-1236(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   TISSUE=Pollen;
RX   MEDLINE=92234570; PubMed=1809687;
RA   Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT   "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT   in Ambrosia artemisiifolia (short ragweed).";
RL   Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M62981; AAA32666.1; -.
DR   EMBL; M80559; AAA32667.1; -.
DR   PIR; B39099; B39099.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    398       POLLEN ALLERGEN AMB A 1.2.
FT   VARIANT     345    345       R -> K.
FT   VARIANT     381    381       L -> I.
SQ   SEQUENCE   398 AA;  43664 MW;  020DC662D9B7416C CRC64;

  Query Match           13.1%;  Score 93;  DB 1;  Length 398;
  Best Local Similarity   30.5%;  Pred. No. 0.022;
  Matches   32;  Conservative    5;  Mismatches   34;  Indels   34;  Gaps    3;

Qy      52 WKNNRIWLQ-----FAKLT-------------------------GFTLMGRRPLWIIF 79
           | |||  |      |||  |                         |    |||||||
Db      62 WANNRQALADCAQGFAKGTYGGKHGDVYTVTSDKDDDVANPKEGTLRFAAAQNRPLWIIF 121

Qy      80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
             || | |   : :   || |||  :|:   ||    :  |  ||
Db     122 KRNMVIHLNQELVVNSDKTIDGRGVKVNI--VNAGLTLMNVKNII 164
```

```
RESULT 11
MPA2_AMBAR
ID   MPA2_AMBAR     STANDARD;      PRT;   397 AA.
AC   P27762;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pollen allergen Amb a 2 precursor (Antigen K) (Antigen Amb a II).
OS   Ambrosia artemisiifolia (Short ragweed).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC   Heliantheae; Ambrosia.
OX   NCBI_TaxID=4212;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Flower;
RX   MEDLINE=92013060; PubMed=1717566;
RA   Rogers B.L., Morgenstern J.P., Griffith I.J., Yu X.-B.,
RA   Counsell C.M., Brauer A.W., King T.P., Garman R.D., Kuo M.-C.C.;
RT   "Complete sequence of the allergen Amb alpha II. Recombinant
RT   expression and reactivity with T cells from ragweed allergic
RT   patients.";
RL   J. Immunol. 147:2547-2552(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   TISSUE=Pollen;
RX   MEDLINE=92234570; PubMed=1809687;
RA   Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT   "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT   in Ambrosia artemisiifolia (short ragweed).";
RL   Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80561; AAA32671.1; -.
DR   PIR; A46469; A46469.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Antigen; Allergen; Signal; Polymorphism.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    397       POLLEN ALLERGEN AMB A 2.
FT   VARIANT      70     70       N -> D (DETECTED ONLY IN FLOWER DNA).
FT   VARIANT     138    138       K -> T (DETECTED ONLY IN FLOWER DNA).
FT   VARIANT     321    321       K -> R (DETECTED ONLY IN FLOWER DNA).
SQ   SEQUENCE   397 AA;  44082 MW;  C78617E4C9A3D1DD CRC64;

  Query Match           12.9%;  Score 92;  DB 1;  Length 397;
  Best Local Similarity   42.3%;  Pred. No. 0.028;
  Matches   22;  Conservative    7;  Mismatches   21;  Indels    2;  Gaps    1;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||  :| | |:  | :   || ||| |:|  |:  |   :  |  :|
Db     114 RPLWIIFQRDMIIYLQQEMVVTSDKTIDGRGAKVELVY--GGITLMNVKNVI 163


RESULT 12
PE56_LYCES
ID   PE56_LYCES     STANDARD;      PRT;   398 AA.
AC   P15721;
```

```
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Probable pectate lyase P56 precursor (EC 4.2.2.2).
GN   LAT56.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VF36; TISSUE=Anther;
RX   MEDLINE=91322485; PubMed=1983191;
RA   Wing R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;
RT   "Molecular and genetic characterization of two pollen-expressed genes
RT   that have sequence similarity to pectate lyases of the plant pathogen
RT   Erwinia.";
RL   Plant Mol. Biol. 14:17-28(1990).
RN   [2]
RP   REVISIONS.
RC   STRAIN=cv. VF36; TISSUE=Anther;
RA   Wing R.A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
CC       GROWTH.
CC   -!- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
CC       oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
CC       their non-reducing ends.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15500; CAA33524.1; -.
DR   PIR; S08383; S08383.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Lyase; Multigene family; Signal.
FT   SIGNAL       1     27       OR 22 (POTENTIAL).
FT   CHAIN       28    398       PROBABLE PECTATE LYASE P56.
FT   ACT_SITE   273    273       POTENTIAL.
FT   CARBOHYD   135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   228    228       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   398 AA;  44563 MW;  8D676250BD8BC7C8 CRC64;

  Query Match             12.8%;  Score 91;  DB 1;  Length 398;
  Best Local Similarity   40.8%;  Pred. No. 0.036;
  Matches   20;  Conservative   10;  Mismatches   15;  Indels    4;  Gaps    2;

Qy      67 FTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV-NGA 114
           | :: : |||| |: :| |:|   : ::  || |||     |||: |||
Db     106 FGVIQKGPLWITFARSMRIRLTRELIVSSNKTIDGRG---KYVHIANGA 151


RESULT 13
MP14_AMBAR
ID   MP14_AMBAR     STANDARD;      PRT;   392 AA.
AC   P28744;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pollen allergen Amb a 1.4 precursor (Antigen E) (Antigen Amb a I).
OS   Ambrosia artemisiifolia (Short ragweed).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC   Heliantheae; Ambrosia.
OX   NCBI_TaxID=4212;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   TISSUE=Pollen;
RX   MEDLINE=92234570; PubMed=1809687;
RA   Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT   "Sequence polymorphism of Amb a I and Amb a II, the major allergens
RT   in Ambrosia artemisiifolia (short ragweed).";
RL   Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC   -!- PTM: The N-terminus is blocked.
CC   -!- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC       AMB A I/AMB A II/CRY J I SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80562; AAA32670.1; -.
DR   PIR; D53240; D53240.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT   SIGNAL       1     25       POTENTIAL.
FT   CHAIN       26    392       POLLEN ALLERGEN AMB A 1.4.
FT   VARIANT    182    188       SHDGPPV -> CNDGPPA.
SQ   SEQUENCE   392 AA;  42842 MW;  7B6219C12F365DA9 CRC64;

  Query Match             12.5%;  Score 89;  DB 1;  Length 392;
  Best Local Similarity   44.2%;  Pred. No. 0.058;
  Matches   23;  Conservative    7;  Mismatches   20;  Indels    2;  Gaps    1;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||: :| |:|   : |   || ||| |:|  ::   |  |  |  ||
Db     114 RPLWIIFARDMVIRLDRELAINNDKTIDGRGAKVEIINAGFA--IYNVKNII 163


RESULT 14
PEL_LILLO
ID   PEL_LILLO      STANDARD;      PRT;   434 AA.
AC   P40973;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Pectate lyase precursor (EC 4.2.2.2).
OS   Lilium longiflorum (Trumpet lily).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
OC   Lilium.
OX   NCBI_TaxID=4690;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nellie white; TISSUE=Pollen;
RA   Kim S.R., Finkel D.J., An G.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
CC       oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
CC       their non-reducing ends.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; Z17328; CAA78976.1; -.
DR   EMBL; L18911; AAA33398.1; -.
DR   PIR; S29612; S29612.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Lyase; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    434       PECTATE LYASE.
FT   ACT_SITE    312    312       POTENTIAL.
FT   CARBOHYD     68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   434 AA;  48457 MW;  C1F3E30AD2BBD064 CRC64;

  Query Match             12.5%;  Score 89;  DB 1;  Length 434;
  Best Local Similarity   55.9%;  Pred. No. 0.064;
  Matches   19;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEV 106
           |||||||  :| |:||  : |   || ||| | |
Db     151 RPLWIIFGKSMVIRLKQELIINNDKTIDGRGANV 184


RESULT 15
PE59_LYCES
ID   PE59_LYCES     STANDARD;      PRT;   449 AA.
AC   P15722;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Probable pectate lyase P59 precursor (EC 4.2.2.2).
GN   LAT59.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VF36; TISSUE=Anther;
RX   MEDLINE=91322485; PubMed=1983191;
RA   Wing R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;
RT   "Molecular and genetic characterization of two pollen-expressed genes
RT   that have sequence similarity to pectate lyases of the plant pathogen
RT   Erwinia.";
RL   Plant Mol. Biol. 14:17-28(1990).
CC   -!- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
CC       GROWTH.
CC   -!- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
CC       oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
CC       their non-reducing ends.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
CC   -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC   ----------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; X15499; CAA33523.1; -.
DR   PIR; S27098; S27098.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Lyase; Multigene family; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    449       PROBABLE PECTATE LYASE P59.
FT   ACT_SITE    325    325       POTENTIAL.
FT   CARBOHYD     56     56       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     80     80       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     81     81       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   449 AA;  50893 MW;  17E3AA13F173B03C CRC64;

  Query Match             11.6%;  Score 82.5;  DB 1;  Length 449;
  Best Local Similarity   22.3%;  Pred. No. 0.32;
  Matches   39;  Conservative   13;  Mismatches   48;  Indels   75;  Gaps    7;

Qy       3 VTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRR---------- 51
           ||:||||     |  :| :  |             |  |      | | :
Db      57 VTLAFNQ---KLRDTMKELKKV-----------KGTHNNSTRRGLGTKKYTGPCMVTNP 101

Qy      52 ----WKNNRIWLQFAKLTGFTLMG---------------------------------- 71
               |: :  |   |      ||
Db     102 IDKCWRCDPNWADNRKKLADCAMGFGSKAIGGKDGEFYVVTDNSDDYNDPKPGTLRHAVI 161

Qy      72 -RRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV-NGA----KFIRRV 120
            : ||||||   |||:|   | :   || | |      ||:  ||    ::|: |
Db     162 QKEPLWIIFKRGMNIRLHQEMIMQSDKTIDARGVN---VHITKGAGITLQYIKNV 213


Search completed: May  5, 2003, 16:57:32
Job time : 21 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 5, 2003, 16:54:26 ; Search time 84 Seconds
(without alignments)
328.694 Million cell updates/sec

Title: US-09-142-524D-3
Perfect score: 711
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....KFIRRVDGIIAAYQNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	188	26.4	374	10	Q8RUR1	Q8rur1 cryptomeria
2	150	21.1	367	10	Q9LLT2	Q9llt2 juniperus v
3	150	21.1	367	10	Q9LLT1	Q9llt1 juniperus v
4	150	21.1	367	10	Q9M4S6	Q9m4s6 cupressus s
5	150	21.1	367	10	Q9M4S4	Q9m4s4 cupressus s
6	150	21.1	367	10	Q9M4S3	Q9m4s3 cupressus s
7	150	21.1	367	10	Q9M4S2	Q9m4s2 cupressus s
8	150	21.1	367	10	Q93X51	Q93x51 juniperus o
9	146	20.5	367	10	Q9M4S5	Q9m4s5 cupressus s
10	144	20.3	367	10	Q93XL6	Q93xl6 cupressus a
11	110.5	15.5	501	10	Q93Z04	Q93z04 arabidopsis
12	110.5	15.5	542	10	Q9SV40	Q9sv40 arabidopsis
13	107.5	15.1	507	10	Q9FY19	Q9fy19 juniperus a
14	104.5	14.7	463	10	Q9SCP2	Q9scp2 arabidopsis
15	100.5	14.1	453	10	Q94LR5	Q94lr5 oryza sativ
16	99	13.9	472	10	Q9LFP5	Q9lfp5 arabidopsis
17	97	13.6	455	10	O64510	O64510 arabidopsis
18	95.5	13.4	181	10	O23666	O23666 arabidopsis
19	95.5	13.4	410	10	Q9FMK5	Q9fmk5 arabidopsis
20	95.5	13.4	432	10	Q93Z25	Q93z25 arabidopsis
21	95.5	13.4	459	10	O23665	O23665 arabidopsis
22	95.5	13.4	459	10	Q9M9S2	Q9m9s2 arabidopsis
23	95	13.4	392	10	Q9FM66	Q9fm66 arabidopsis
24	95	13.4	452	10	Q9LRM5	Q9lrm5 arabidopsis
25	94.5	13.3	226	10	O23667	O23667 arabidopsis
26	94.5	13.3	354	10	Q940Q1	Q940q1 arabidopsis
27	94.5	13.3	398	10	Q9M505	Q9m505 vitis vinif
28	94.5	13.3	404	10	Q9SB71	Q9sb71 arabidopsis
29	94.5	13.3	408	10	Q9C5M8	Q9c5m8 arabidopsis
30	94.5	13.3	431	10	O23017	O23017 arabidopsis
31	93.5	13.2	392	10	Q9FXD8	Q9fxd8 arabidopsis
32	93.5	13.2	407	10	Q9SDW4	Q9sdw4 musa acumin
33	93.5	13.2	408	10	Q8W116	Q8w116 arabidopsis
34	93	13.1	450	10	Q40319	Q40319 medicago sa
35	91.5	12.9	368	10	Q9C8G4	Q9c8g4 arabidopsis
36	90.5	12.7	409	10	Q93XJ1	Q93xj1 salix gilgi
37	90	12.7	409	10	Q9LTZ0	Q9ltz0 arabidopsis
38	89.5	12.6	324	10	Q8S345	Q8s345 capsicum an
39	89.5	12.6	398	10	Q43783	Q43783 musa acumin
40	88.5	12.4	297	10	Q9FY87	Q9fy87 arabidopsis
41	88	12.4	438	10	Q43862	Q43862 zea mays (m
42	87	12.2	368	10	Q94FT5	Q94ft5 fragaria an
43	85.5	12.0	416	10	Q9M8Z8	Q9m8z8 arabidopsis
44	85	12.0	227	10	O23668	O23668 arabidopsis
45	84.5	11.9	401	10	O24554	O24554 zinnia eleg

ALIGNMENTS

RESULT 1
Q8RUR1
ID Q8RUR1 PRELIMINARY; PRT; 374 AA.
AC Q8RUR1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cry j 1 precursor.
GN CRY J 1.1 OR CRY J 1.2.
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Futamura N., Shinohara K.;
RT "Isolation and characterization of cDNAs encoding major allergen Cry j
RT 1 from Cryptomeria japonica pollen.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB081309; BAB86286.1; -.
DR EMBL; AB081310; BAB86287.1; -.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 374 CRY J 1.
SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 26.4%; Score 188; DB 10; Length 374;
Best Local Similarity 71.2%; Pred. No. 2.6e-12;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138
```

RESULT 2
Q9LLT2
ID Q9LLT2 PRELIMINARY; PRT; 367 AA.

```
AC   Q9LLT2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pollen major allergen 1-2.
OS   Juniperus virginiana (Eastern red cedar).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX   NCBI_TaxID=39584;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21315424; PubMed=11422137;
RA   Midoro-Horiuti T.M., Goldblum R.M., Brooks E.G.;
RT   "Identification of mutations in the genes for the pollen allergens of
RT   eastern red cedar (Juniperus virginiana).";
RL   Clin. Exp. Allergy 31:771-778(2001).
DR   EMBL; AF151427; AAF80164.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39768 MW;  0A6AC2F1BAF89586 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||||:|:||:|| ||| |:|   :      |:|:|  :|
Db      87 KTLWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 3
Q9LLT1
ID   Q9LLT1          PRELIMINARY;      PRT;   367 AA.
AC   Q9LLT1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pollen major allergen 1-1.
OS   Juniperus virginiana (Eastern red cedar).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX   NCBI_TaxID=39584;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21315424; PubMed=11422137;
RA   Midoro-Horiuti T.M., Goldblum R.M., Brooks E.G.;
RT   "Identification of mutations in the genes for the pollen allergens of
RT   eastern red cedar (Juniperus virginiana).";
RL   Clin. Exp. Allergy 31:771-778(2001).
DR   EMBL; AF151429; AAF80166.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39708 MW;  DCBD1981A74E4711 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||||:|:||:|| ||| |:|   :      |:|:|  :|
Db      87 KTLWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 4
Q9M4S6
ID   Q9M4S6      PRELIMINARY;      PRT;   367 AA.
AC   Q9M4S6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cup s 1 pollen allergen.
GN   CUPS1.
OS   Cupressus sempervirens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=13469;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Monsalve R.I., Villalba M., Rodriguez R.;
RT   "Cloning and expression of Cup s 1, the major allergen of the pollen
RT   of Cupressus sempervirens.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257491; AAF72625.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39835 MW;  B4E9C60108C2C5A3 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||||:|:||:|| ||| |:|   :      |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 5
Q9M4S4
ID   Q9M4S4      PRELIMINARY;      PRT;   367 AA.
AC   Q9M4S4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cup s 1 pollen allergen.
GN   CUPS1.
OS   Cupressus sempervirens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=13469;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Monsalve R.I., Villalba M., Rodriguez R.;
RT   "Cloning and expression of Cup s 1, the major allergen of the pollen
RT   of Cupressus sempervirens.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257493; AAF72627.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39934 MW;  974D3011D74E3D6E CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||||:|:||:|| ||| |:|   :      |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 6
Q9M4S3
ID   Q9M4S3      PRELIMINARY;      PRT;   367 AA.
AC   Q9M4S3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cup s 1 pollen allergen.
GN   CUPS1.
OS   Cupressus sempervirens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=13469;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Monsalve R.I., Villalba M., Rodriguez R.;
RT   "Cloning and expression of Cup s 1, the major allergen of the pollen
RT   of Cupressus sempervirens.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257494; AAF72628.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39832 MW;  B5DFBF5A61C07A53 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| ||||||||:|:||:|| ||| |:|   :     |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 7
Q9M4S2
ID   Q9M4S2      PRELIMINARY;      PRT;  367 AA.
AC   Q9M4S2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cup s 1 pollen allergen.
GN   CUPS1.
OS   Cupressus sempervirens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=13469;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Monsalve R.I., Villalba M., Rodriguez R.;
RT   "Cloning and expression of Cup s 1, the major allergen of the pollen
RT   of Cupressus sempervirens.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257495; AAF72629.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39819 MW;  AE7E055A61C07A53 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| ||||||||:|:||:|| ||| |:|   :     |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 8
Q93X51
ID   Q93X51      PRELIMINARY;      PRT;  367 AA.
AC   Q93X51;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative allergen jun o 1.
GN   JUN O 1.
OS   Juniperus oxycedrus (Prickly juniper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX   NCBI_TaxID=69008;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=POLLEN;
RA   Iacovacci P., Di Felice G., Pini C.;
RT   "Cloning of Juniperus oxycedrus major allergen.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ293767; CAC48400.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
SQ   SEQUENCE   367 AA;  39808 MW;  5D28204DBFD1B9D7 CRC64;

  Query Match             21.1%;  Score 150;  DB 10;  Length 367;
  Best Local Similarity   55.8%;  Pred. No. 3.7e-08;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| ||||||||:|:||:|| ||| |:|   :     |:|:|  :|
Db      87 KALWIIFSQNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 9
Q9M4S5
ID   Q9M4S5      PRELIMINARY;      PRT;  367 AA.
AC   Q9M4S5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cup s 1 pollen allergen.
GN   CUPS1.
OS   Cupressus sempervirens.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=13469;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Monsalve R.I., Villalba M., Rodriguez R.;
RT   "Cloning and expression of Cup s 1, the major allergen of the pollen
RT   of Cupressus sempervirens.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257492; AAF72626.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
SQ   SEQUENCE   367 AA;  39894 MW;  5D56FC0E3263B741 CRC64;

  Query Match             20.5%;  Score 146;  DB 10;  Length 367;
  Best Local Similarity   53.8%;  Pred. No. 1e-07;
  Matches   28;  Conservative   10;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| ||||||:||:|:||:|| ||| |:|   :     |:|:|  :|
Db      87 KALWIIFSQNMNIKLEMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSHVI 138


RESULT 10
Q93XL6
ID   Q93XL6      PRELIMINARY;      PRT;  367 AA.
AC   Q93XL6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative allergen Cup a 1 precursor.
GN   CUP A 1.
OS   Cupressus arizonica.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX   NCBI_TaxID=49011;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=POLLEN;
RA   Butteroni C., Di Felice G., Pini C.;
RT   "Cloning of Cupressus Arizonica major allergen.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ278498; CAC37790.2; -.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    367       PUTATIVE ALLERGEN CUP A 1.
SQ   SEQUENCE   367 AA;  39809 MW;  AFF97260423A9F28 CRC64;
```

```
  Query Match           20.3%;  Score 144;  DB 10;  Length 367;
  Best Local Similarity  53.8%;  Pred. No. 1.7e-07;
  Matches   28;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           : |||||| |||||||:||:|:||||| ||| |:|   :     |:|    :|
Db      87 KALWIIFSQNMNIKLQMPLYVAGYKTIDGRGADVHLGNGGPCLFMRTASHVI 138


RESULT 11
Q93Z04
ID   Q93Z04      PRELIMINARY;      PRT;   501 AA.
AC   Q93Z04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   AT3g54920/F28P10_100 (Putative pectate lyase).
GN   AT3G54920.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA   Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA   Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA   Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA   Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA   Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RT   "Full Length cDNA of gene At3g54920 (GI:15233132).";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY058870; AAL24257.1; -.
DR   EMBL; AY074331; AAL67027.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
KW   Lyase.
SQ   SEQUENCE   501 AA;  53929 MW;  9E0DE36DEF4C7ABB CRC64;

  Query Match           15.5%;  Score 110.5;  DB 10;  Length 501;
  Best Local Similarity  43.3%;  Pred. No. 0.0011;
  Matches   26;  Conservative    9;  Mismatches   20;  Indels    5;  Gaps    2;

Qy      67 FTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV--NGAKFIRRVDGII 124
           : ::   ||||:|| || |:||  : |  ||| |||    | ||:  ||   :: |  ||
Db     139 YGVIQEEPLWIVFSSNMLIRLKQELIINSYKTLDGRG---SAVHITGNGCLTLQYVQHII 195


RESULT 12
Q9SV40
ID   Q9SV40      PRELIMINARY;      PRT;   542 AA.
AC   Q9SV40;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Pectate lyase-like protein.
GN   F28P10.100.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Choisne N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA   Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Mayer K.F.X.,
RA   Lemcke K., Schueller C., Quetier F., Salanoubat M.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL049655; CAB41092.1; -.
DR   InterPro; IPR002022; Amb_allergen.
DR   Pfam; PF00544; pec_lyase; 1.
DR   PRINTS; PR00807; AMBALLERGEN.
KW   Lyase.
SQ   SEQUENCE   542 AA;  58573 MW;  EB3779D613B72347 CRC64;

  Query Match           15.5%;  Score 110.5;  DB 10;  Length 542;
  Best Local Similarity  43.3%;  Pred. No. 0.0013;
  Matches   26;  Conservative    9;  Mismatches   20;  Indels    5;  Gaps    2;

Qy      67 FTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV--NGAKFIRRVDGII 124
           : ::   ||||:|| || |:||  : |  ||| |||    | ||:  ||   :: |  ||
Db     139 YGVIQEEPLWIVFSSNMLIRLKQELIINSYKTLDGRG---SAVHITGNGCLTLQYVQHII 195


RESULT 13
Q9FY19
ID   Q9FY19      PRELIMINARY;      PRT;   507 AA.
AC   Q9FY19;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pollen major allergen 2 protein precursor.
GN   JNA2.
OS   Juniperus ashei (Ozark white cedar).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX   NCBI_TaxID=13101;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MALE POLLEN;
RX   MEDLINE=20403896; PubMed=10944464;
RA   Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;
RT   "Purification, Identification and cDNA cloning of Jun a 2, the second
RT   major allergen of mountain cedar pollen.";
RL   Biochem. Biophys. Res. Commun. 275:195-202(2000).
CC   -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC       (POLYGALACTURONASES).
DR   EMBL; AJ404653; CAC05582.1; -.
DR   HSSP; P26509; 1BHE.
DR   InterPro; IPR000743; GH28.
DR   InterPro; IPR000408; Reg_chr_condens.
DR   Pfam; PF00295; Glyco_hydro_28; 1.
DR   PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
DR   PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW   Cell wall; Glycosidase; Hydrolase; Signal.
FT   SIGNAL       1     54       POTENTIAL.
SQ   SEQUENCE   507 AA;  55730 MW;  2B2E0AA5E958FE5A CRC64;

  Query Match           15.1%;  Score 107.5;  DB 10;  Length 507;
  Best Local Similarity  32.3%;  Pred. No. 0.0025;
  Matches   32;  Conservative   13;  Mismatches   25;  Indels   29;  Gaps    2;

Qy      31 IDIFASKNFHLQKNTIGTGRRWKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
           ||||||| | ::| |||||                              : :|: || :|
Db     237 IDIFASKRFEIEKCTIGTG----------------------DDCVAVGTGSSNITIKDL 273
```

```
Qy      91 MYIAGYKTFDG------RRAEVSYVHVNGAKFIRRVDGI 123
           |:    |       |:|||:||::|||||   :|:
Db     274 TCGPGHGMSIGSLGKGNSRSEVSFVHLDGAKFIDTQNGL 312
```

RESULT 14
Q9SCP2
ID Q9SCP2 PRELIMINARY; PRT; 463 AA.
AC Q9SCP2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Pectate lyase-like protein.
GN T4D2.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL132958; CAB64222.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Lyase.
SQ SEQUENCE 463 AA; 51225 MW; 3424B0DE4ADBA604 CRC64;

  Query Match 14.7%; Score 104.5; DB 10; Length 463;
  Best Local Similarity 30.5%; Pred. No. 0.0047;
  Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;

```
Qy      52 WKNNRIWL----------------QFAKLTG--------------FTLMGRRPLWIIF 79
           | |||  |                ||  :|               : ::   ||||:|
Db      66 WPNNRQGLADCGIGFGQYALGGKGGQFYFVTDSSDDDAVNPKPGTLRYGVIQEEPLWIVF 125

Qy      80 SGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
             || ||||  :    ||| ||| | |  |   |   :: |  ||
Db     126 PSNMMIKLKQELIFNSYKTLDGRGANVHIVG-GGCITLQYVSNII 169
```

RESULT 15
Q94LR5
ID Q94LR5 PRELIMINARY; PRT; 453 AA.
AC Q94LR5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative pectate lyase.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBb0011A08 genomic sequence.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC034258; AAK54283.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
KW Lyase.
SQ SEQUENCE 453 AA; 48006 MW; 1411BEB1A40901DC CRC64;

  Query Match 14.1%; Score 100.5; DB 10; Length 453;
  Best Local Similarity 48.8%; Pred. No. 0.013;
  Matches 20; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

```
Qy      74 PLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGA 114
           ||||:|:|:| |:|   : :  ||| ||| | | :|   ||
Db     140 PLWIVFAGDMTIRLNEELLVNSYKTIDGRGANV-HVGAGGA 179
```

Search completed: May 5, 2003, 16:59:01
Job time : 90 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 5, 2003, 16:48:06 ; Search time 46 Seconds
(without alignments)
388.165 Million cell updates/sec

Title: US-09-142-524D-3
Perfect score: 711
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....KFIRRVDGIIAAYQNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	711	100.0	134	18	AAW27371	Multi-epitope pept
2	523.5	73.6	105	18	AAW27370	Multi-epitope pept
3	320	45.0	80	18	AAW27369	Multi-epitope pept
4	198.5	27.9	214	22	AAB69120	Cedar pollen aller
5	188	26.4	353	16	AAR75388	Japanese cedar pol
6	188	26.4	353	17	AAR81587	Cedar pollen aller
7	188	26.4	373	20	AAY25664	Cedar allergen 493
8	188	26.4	373	20	AAY25668	Japanese cedar all
9	188	26.4	374	14	AAR31937	Cry j I. Cryptome
10	188	26.4	374	15	AAR45541	Cry j I pollen all
11	188	26.4	374	15	AAR60166	Japanese cedar pol
12	188	26.4	374	16	AAR82490	Cry j I Japanese C
13	188	26.4	374	20	AAY25665	Cedar allergen 493
14	188	26.4	374	20	AAY25669	Japanese cedar all
15	178	25.0	210	22	AAB69103	Cedar pollen aller
16	162	22.8	96	22	AAB69104	Cedar pollen aller
17	162	22.8	96	22	AAB69119	Cedar pollen aller
18	162	22.8	97	22	AAB69105	Cedar pollen aller
19	150	21.1	367	15	AAR45577	Jun s I. Juniperu
20	150	21.1	370	15	AAR45578	Jun v I. Juniperu
21	148.5	20.9	460	16	AAR69791	Japonicum allergen
22	148.5	20.9	514	15	AAR53690	Japanese cedar pol
23	148.5	20.9	514	16	AAR74333	Japanese cedar pol
24	148.5	20.9	514	16	AAR69792	Japonicum allergen
25	148.5	20.9	514	17	AAR93599	Japan cedar pollen
26	148.5	20.9	514	17	AAR81586	Cedar pollen aller
27	148.5	20.9	514	20	AAY25666	Japanese cedar all
28	148.5	20.9	514	20	AAY25667	Japanese cedar all
29	148	20.8	186	22	AAB69102	Cedar pollen aller
30	142	20.0	354	17	AAW04344	Chamaecyparis obtu
31	142	20.0	354	19	AAW42121	Japanese cypress p
32	142	20.0	375	17	AAW04345	Chamaecyparis obtu
33	141.5	19.9	93	22	AAB69117	Cedar pollen aller
34	136	19.1	30	19	AAW44682	T-cell epitope pep
35	120.5	16.9	74	21	AAB23905	Artificial sequenc
36	120.5	16.9	94	21	AAB23906	Artificial sequenc
37	120.5	16.9	95	21	AAB23897	Artificial sequenc
38	120.5	16.9	95	21	AAB23901	Artificial sequenc
39	120.5	16.9	95	22	AAB69098	Cedar pollen aller
40	120.5	16.9	99	21	AAB23902	Artificial sequenc
41	120.5	16.9	99	21	AAB23907	Artificial sequenc
42	120.5	16.9	99	22	AAB69099	Cedar pollen aller
43	118.5	16.7	514	17	AAW04346	Chamaecyparis obtu
44	118.5	16.7	514	19	AAW42122	Japanese cypress p
45	117	16.5	47	19	AAW80353	Sugi allergen prot

ALIGNMENTS

RESULT 1
AAW27371
ID AAW27371 standard; peptide; 134 AA.
XX
AC AAW27371;
XX
DT 24-MAR-1998 (first entry)
XX
DE Multi-epitope peptide used as immunotherapeutic agent #3.
XX
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
XX
OS Synthetic.
XX
PN WO9732600-A1.
XX
PD 12-SEP-1997.
XX
PF 10-MAR-1997; 97WO-JP00740.
XX
PR 10-MAR-1996; 96JP-0080702.
XX
PA (MEIP) MEIJI MILK PROD CO LTD.
XX
PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;
XX
DR WPI; 1997-470495/43.
XX
PT Peptide immuno:therapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
XX

```
PS   Claim 6; Page 32; 58pp; Japanese.
XX
CC   The present sequence represents a multi-epitope peptide which is used as
CC   a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC   or more different allergens (preferably linked via arginine or lysine
CC   dimers), where the T cell epitope regions: have a positivity index
CC   greater than 100 as measured in a patient group responding to the
CC   allergen; have at least 70% reactivity with lymphocytes from patients
CC   responding to the allergen; and are not reactive with immunoglobulin E
CC   (IgE) antibodies from patients responsive to the allergen. The agent can
CC   be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC   desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
XX
SQ   Sequence   134 AA;

  Query Match             100.0%;  Score 711;  DB 18;  Length 134;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-82;
  Matches  134;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQ 60

Qy      61 FAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 FAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRV 120

Qy     121 DGIIAAYQNPASWK 134
           ||||||||||||||
Db     121 DGIIAAYQNPASWK 134


RESULT 2
AAW27370
ID   AAW27370 standard; peptide; 105 AA.
XX
AC   AAW27370;
XX
DT   24-MAR-1998  (first entry)
XX
DE   Multi-epitope peptide used as immunotherapeutic agent #2.
XX
KW   Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW   T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
XX
OS   Synthetic.
XX
PN   WO9732600-A1.
XX
PD   12-SEP-1997.
XX
PF   10-MAR-1997;   97WO-JP00740.
XX
PR   10-MAR-1996;   96JP-0080702.
XX
PA   (MEIP ) MEIJI MILK PROD CO LTD.
XX
PI   Dairiki K,  Iwama A,  Kino K,  Kume A,  Sone T;
XX
DR   WPI; 1997-470495/43.
XX
PT   Peptide immuno:therapeutic agent to treat allergic diseases -
PT   contains multi-epitope peptide containing T cell epitope regions
PT   from different allergens
XX
PS   Claim 6; Page 31; 58pp; Japanese.
XX
CC   The present sequence represents a multi-epitope peptide which is used as
CC   a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC   or more different allergens (preferably linked via arginine or lysine
CC   dimers), where the T cell epitope regions: have a positivity index
CC   greater than 100 as measured in a patient group responding to the
CC   allergen; have at least 70% reactivity with lymphocytes from patients
CC   responding to the allergen; and are not reactive with immunoglobulin E
CC   (IgE) antibodies from patients responsive to the allergen. The agent can
CC   be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC   desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
XX
SQ   Sequence   105 AA;

  Query Match             73.6%;  Score 523.5;  DB 18;  Length 105;
  Best Local Similarity   78.4%;  Pred. No. 9.8e-59;
  Matches  105;  Conservative    0;  Mismatches    0;  Indels   29;  Gaps    2;

Qy       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQ 60

Qy      61 FAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRV 120
           ||||||||||||          ||||||||||||||                  |||
Db      61 FAKLTGFTLMGRR----------LKMPMYIAGYKTFDG----------------RRV 91

Qy     121 DGIIAAYQNPASWK 134
           ||||||||||||||
Db      92 DGIIAAYQNPASWK 105


RESULT 3
AAW27369
ID   AAW27369 standard; peptide; 80 AA.
XX
AC   AAW27369;
XX
DT   24-MAR-1998  (first entry)
XX
DE   Multi-epitope peptide used as immunotherapeutic agent #1.
XX
KW   Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW   T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
XX
OS   Synthetic.
XX
PN   WO9732600-A1.
XX
PD   12-SEP-1997.
XX
PF   10-MAR-1997;   97WO-JP00740.
XX
PR   10-MAR-1996;   96JP-0080702.
XX
PA   (MEIP ) MEIJI MILK PROD CO LTD.
XX
PI   Dairiki K,  Iwama A,  Kino K,  Kume A,  Sone T;
XX
DR   WPI; 1997-470495/43.
XX
PT   Peptide immuno:therapeutic agent to treat allergic diseases -
PT   contains multi-epitope peptide containing T cell epitope regions
PT   from different allergens
XX
PS   Claim 6; Page 31; 58pp; Japanese.
XX
CC   The present sequence represents a multi-epitope peptide which is used as
CC   a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC   or more different allergens (preferably linked via arginine or lysine
CC   dimers), where the T cell epitope regions: have a positivity index
CC   greater than 100 as measured in a patient group responding to the
CC   allergen; have at least 70% reactivity with lymphocytes from patients
CC   responding to the allergen; and are not reactive with immunoglobulin E
CC   (IgE) antibodies from patients responsive to the allergen. The agent can
CC   be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC   desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
XX
SQ   Sequence   80 AA;

  Query Match             45.0%;  Score 320;  DB 18;  Length 80;
```

```
  Best Local Similarity   53.7%;  Pred. No. 6.1e-33;
  Matches   72;  Conservative    2;  Mismatches    6;  Indels   54;  Gaps    2;

Qy       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRR--------- 51

Qy      61 FAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRV 120
                              | ||:                    :|    |||
Db      52 ----------------------ISLKL--------------------TSGKIASRRV 66

Qy     121 DGIIAAYQNPASWK 134
           ||||||||||||||
Db      67 DGIIAAYQNPASWK 80


RESULT 4
AAB69120
ID   AAB69120 standard; Protein; 214 AA.
XX
AC   AAB69120;
XX
DT   23-APR-2001  (first entry)
XX
DE   Cedar pollen allergen T cell epitope derived protein SEQ ID NO:61.
XX
KW   Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
KW   T cell epitope; antisugipollinosis.
XX
OS   Cryptomeria japonica.
OS   Synthetic.
XX
PN   JP2000327699-A.
XX
PD   28-NOV-2000.
XX
PF   15-MAR-2000; 2000JP-0071710.
XX
PR   15-MAR-1999;   99JP-0068316.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
PA   (SANY ) SANKYO CO LTD.
XX
DR   WPI; 2001-185061/19.
DR   N-PSDB; AAF59044.
XX
PT   Novel peptide and its use  -
XX
PS   Example 11; Page 58-59; 75pp; Japanese.
XX
CC   The present invention describes a peptide, its complex, derivative or
CC   its polymerizate, where the peptide (I) has a formula of:
CC   alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
CC   where alpha1 to alpha7 = amino acid sequence selected from amino acid
CC   sequences ranging from 11-19 amino acids derived from T cell epitopes
CC   derived from cedar (japanese cedar - Cryptomeria japonica) pollen
CC   allergens. The peptide can be used in an antisugipollinosis agent.
CC   AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
CC   in the exemplification of the present invention.
XX
SQ   Sequence   214 AA;

  Query Match             27.9%;  Score 198.5;  DB 22;  Length 214;
  Best Local Similarity   33.7%;  Pred. No. 7.4e-17;
  Matches   55;  Conservative    7;  Mismatches    8;  Indels   93;  Gaps    5;

Qy       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGR---------- 50
           ||||||||||||                    ||||||||||||| : .:|:
Db      30 MKVTVAFNQFGP------------------DIFASKNFHLQKNKLTSGKIASCLNYGLV 70

Qy      51 -----------RWKNNRIWL--------QFAKLTGFTLMGRRPLWIIFSGNMNIKLKMPM 91
                     :::   |:         ||||||||||||
Db      71 HVANNNYDPSGKYEGGNIYTKKEAFNVEQFAKLTGFTLMGRA------------------ 112

Qy      92 YIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGIIAAYQNPASWK 134
                 | |                    ||||||||||||
Db     113 --------DPR-------------------GIIAAYQNPASWK 128


RESULT 5
AAR75388
ID   AAR75388 standard; protein; 353 AA.
XX
AC   AAR75388;
XX
DT   12-MAR-1996  (first entry)
XX
DE   Japanese cedar pollen allergen Cryj I.
XX
KW   Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
KW   prevention; treatment; cryptomeria pollinosis.
XX
OS   Cryptomeria japonica.
XX
FH   Key             Location/Qualifiers
FT   Peptide         61..75
FT                   /note= "T-cell epitope peptide"
FT   Peptide         91..105
FT                   /note= "T-cell epitope peptide"
FT   Peptide         106..120
FT                   /note= "T-cell epitope peptide"
FT   Peptide         146..160
FT                   /note= "T-cell epitope peptide"
FT   Peptide         211..225
FT                   /note= "T-cell epitope peptide"
FT   Peptide         326..340
FT                   /note= "T-cell epitope peptide"
FT   Peptide         335..346
FT                   /note= "T-cell epitope peptide"
XX
PN   JP07118295-A.
XX
PD   09-MAY-1995.
XX
PF   20-OCT-1993;   93JP-0262626.
XX
PR   20-OCT-1993;   93JP-0262626.
XX
PA   (MEIP ) MEIJI MILK PROD CO LTD.
XX
DR   WPI; 1995-203834/27.
XX
PT   New cryptomeria pollen allergen T-cell epitope peptide - used for
PT   prevention, treatment and investigation of Japanese cedar pollenosis
XX
PS   Disclosure; Figs 1-2; 8pp; Japanese.
XX
CC   AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC   T-cell epitope peptides AAR89289-R89295 were derived. The peptides
CC   can be used for the prevention and treatment of cryptomeria
CC   pollenosis, and also for the investigation of pollinosis.
XX
SQ   Sequence   353 AA;

  Query Match             26.4%;  Score 188;  DB 16;  Length 353;
  Best Local Similarity   71.2%;  Pred. No. 3.2e-15;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      66 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 117


RESULT 6
AAR81587
ID   AAR81587 standard; Protein; 353 AA.
```

XX
AC AAR81587;
XX
DT 24-MAY-1996 (first entry)
XX
DE Cedar pollen allergen B.
XX
KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
KW antibody; pollinosis; therapy; immunotherapy.
XX
OS Cryptomeria japonica.
XX
PN EP700929-A2.
XX
PD 13-MAR-1996.
XX
PF 08-SEP-1995; 95EP-0306295.
XX
PR 14-JUL-1995; 95JP-0200221.
PR 10-SEP-1994; 94JP-0242137.
PR 14-JUL-1995; 95JP-0200204.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hino K, Saito S, Taniguchi Y;
XX
DR WPI; 1996-140976/15.
XX
PT New peptide(s) derived from cedar pollen allergens - activate
PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
PT used for treating cedar pollinosis
XX
PS Claim 5; Page 31-32; 36pp; English.
XX
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (AAR81586) and B (AAR81587) were tested for their ability to activate
CC cedar allergen-specific T-cells, but not allergen-specific IgE
CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
CC for T-cell recognition, and homologous peptides (AAR81588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.
XX
SQ Sequence 353 AA;

Query Match 26.4%; Score 188; DB 17; Length 353;
Best Local Similarity 71.2%; Pred. No. 3.2e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      66 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 117
```

RESULT 7
AAY25664
ID AAY25664 standard; protein; 373 AA.
XX
AC AAY25664;
XX
DT 30-SEP-1999 (first entry)
XX
DE Cedar allergen 493634 Cry j IB protein fragment.
XX
KW Major histocompatibility complex; class II; desensitising; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
OS Cedrus sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI; 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 73; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chiromidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.
XX
SQ Sequence 373 AA;

Query Match 26.4%; Score 188; DB 20; Length 373;
Best Local Similarity 71.2%; Pred. No. 3.5e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      86 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 137
```

RESULT 8
AAY25668
ID AAY25668 standard; protein; 373 AA.
XX
AC AAY25668;
XX
DT 30-SEP-1999 (first entry)
XX
DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.
XX
KW Major histocompatibility complex; class II; desensitising; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
OS Cedrus sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.

PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI; 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 75; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chiromidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a japanese cedar (Cedrus sp.) allergen 541803 Cry j I
CC precursor.
XX
SQ Sequence 373 AA;

Query Match 26.4%; Score 188; DB 20; Length 373;
Best Local Similarity 71.2%; Pred. No. 3.5e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           ||||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      86 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 137
```

RESULT 9
AAR31937
ID AAR31937 standard; Protein; 374 AA.
XX
AC AAR31937;
XX
DT 03-JUN-1993 (first entry)
XX
DE Cry j I.
XX
KW Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Protein 22..374
FT /note= "mature Cry j I"
XX
PN WO9301213-A.
XX
PD 21-JAN-1993.
XX
PF 10-JUL-1992; 92WO-US05661.
XX
PR 12-JUL-1991; 91US-0729134.
PR 15-JUL-1991; 91US-0730452.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Bond JF, Griffith IJ, Pollock J;
XX
DR WPI; 1993-045434/05.
DR N-PSDB; AAQ35304.
XX
PT Nucleic acid sequence encoding Cryptomeria japonica allergen -
PT for the diagnosis treatment and prevention of allergic reactions
PT to Japanese cedar pollen
XX
PS Claim 11; Page 42; 69pp; English.
XX
CC Fresh pollen and staminate cone samples were collected from a single
CC Cryptomeria japonica (Japanese cedar) tree. RNA was prepd. and used
CC to synthesise cDNA. The cDNA was subjected to successive rounds of
CC PCR to yield a full length Cry j I clone. Cry j I or an antigenic
CC fragment of it may be used for detecting, treating and preventing an
CC allergic response to Japanese cedar pollen allergen. It is capable of
CC modifying both the B and T cell response to Cry j I and T cell response
CC to a Cry j I antigen.
XX
SQ Sequence 374 AA;

Query Match 26.4%; Score 188; DB 14; Length 374;
Best Local Similarity 71.2%; Pred. No. 3.5e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           ||||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138
```

RESULT 10
AAR45541
ID AAR45541 standard; Protein; 374 AA.
XX
AC AAR45541;
XX
DT 13-JUL-1994 (first entry)
XX
DE Cry j I pollen allergen.
XX
KW Japanese cedar; detection; allergy; treatment; diagnosis;
KW T cell epitope; sensitivity.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Peptide 22..374
FT /note= "mature peptide"
XX
PN WO9401560-A.
XX
PD 20-JAN-1994.
XX
PF 15-JAN-1993; 93WO-US00139.
XX
PR 01-SEP-1992; 92US-0938990.
PR 10-JUL-1992; 92WO-US05661.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
XX
DR WPI; 1994-035066/04.
DR N-PSDB; AAQ55271.
XX
PT Antigens derived from Japanese cedar pollen allergen Cry j I -
PT contain at least two T cell epitope(s), used to treat or diagnose
PT allergy
XX
PS Disclosure; Fig 4; 137pp; English.
XX

```
CC   The sequence is that of the Japanese cedar pollen allergen
CC   Cry j I which contains at least two T cell epitopes. Peptide
CC   antigens derived from it can be used for the treatment and
CC   diagnosis of allergies associated with Japanese cedar pollen.
CC   The peptides have enhanced therapeutic properties but reduced
CC   side effects compared to naturally occurring allergens.
XX
SQ   Sequence   374 AA;

  Query Match            26.4%;  Score 188;  DB 15;  Length 374;
  Best Local Similarity   71.2%;  Pred. No. 3.5e-15;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138


RESULT 11
AAR60166
ID   AAR60166 standard; Protein; 374 AA.
XX
AC   AAR60166;
XX
DT   24-MAR-1995  (first entry)
XX
DE   Japanese cedar pollen antigen CryjI.
XX
KW   Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.
XX
OS   Cryptomeria japonica.
XX
FH   Key             Location/Qualifiers
FT   Peptide          1..21
FT                    /label= signal_peptide
FT   Protein          22..374
FT                    /label= mature_CryjI
XX
PN   JP06197768-A.
XX
PD   19-JUL-1994.
XX
PF   07-JAN-1993;   93JP-0001116.
XX
PR   07-JAN-1993;   93JP-0001116.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA.
XX
DR   WPI; 1994-268680/33.
DR   N-PSDB; AAQ71601.
XX
PT   Sugi (Japanese cedar) pollen antigen CryjI - is  useful for
PT   diagnosis, treatment and prevention of sugi pollinosis
XX
PS   Claim 2; Page 5-7; 9pp; Japanese.
XX
CC   The coding sequence for the Japanese cedar ("sugi") pollen allergen
CC   CryjI was isolated from a cDNA library prepared from polyA mRNA. All
CC   or part of the CryjI protein can be used for diagnosis, treatment
CC   and prevention of sugi pollinosis.
XX
SQ   Sequence   374 AA;

  Query Match            26.4%;  Score 188;  DB 15;  Length 374;
  Best Local Similarity   71.2%;  Pred. No. 3.5e-15;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138


RESULT 12
AAR82490
ID   AAR82490 standard; Protein; 374 AA.
XX
AC   AAR82490;
XX
DT   15-APR-1996  (first entry)
XX
DE   Cry j I Japanese Cedar pollen allergen.
XX
KW   Cry j I; japanese cedar pollen allergen; modified; drug production;
KW   allergy; Crytpomeria japonica.
XX
OS   Crytpomeria japonica.
XX
PN   WO9527786-A1.
XX
PD   19-OCT-1995.
XX
PF   06-APR-1995;   95WO-US04249.
XX
PR   06-DEC-1994;   94US-0350225.
PR   08-APR-1994;   94US-0226248.
XX
PA   (IMMU-) IMMULOGIC PHARM CORP.
XX
PI   Chen X,  Evans S,  Franzen HM,  Kuo M,  Powers SP;
PI   Shaked Z;
XX
DR   WPI; 1995-366391/47.
DR   N-PSDB; AAT04248.
XX
PT   Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for
PT   treating allergy to japanese cedar pollen allergen or
PT   immunologically cross reactive allergens
XX
PS   Disclosure; Figure 1; 60pp; English.
XX
CC   Novel peptides of cry j I have been modified as a part of a
CC   preformulation scheme to develop an optimised drug product for
CC   therapeutic treatment of humans suffering from allergy to Japanese
CC   cedar pollen allergen or an allergen which is immunologically cross
CC   reactive with Japanese cedar pollen allergen. Such modified peptides
CC   possess certain characteristics which render them particularly
CC   suitable for drug product formulation.  Peptide fragments of Cry j I,
CC   modified and unmodified, are given in AAR82491-R82525.
XX
SQ   Sequence   374 AA;

  Query Match            26.4%;  Score 188;  DB 16;  Length 374;
  Best Local Similarity   71.2%;  Pred. No. 3.5e-15;
  Matches   37;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
           |||||||||||||||||||||||||||||| |:|   :     ||:||  :|
Db      87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138


RESULT 13
AAY25665
ID   AAY25665 standard; protein; 374 AA.
XX
AC   AAY25665;
XX
DT   30-SEP-1999  (first entry)
XX
DE   Cedar allergen 493632 Cry j IA protein fragment.
XX
KW   Major histocompatibility complex; class II; desensitising; human;
KW   allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW   chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW   screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW   cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW   mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
```

XX
OS Cedrus sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI; 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 73; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chiromidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.
XX
SQ Sequence 374 AA;

Query Match 26.4%; Score 188; DB 20; Length 374;
Best Local Similarity 71.2%; Pred. No. 3.5e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
||||||||||||||||||||||||||||||||| |:| : ||:|| :|
Db 87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138

RESULT 14
AAY25669
ID AAY25669 standard; protein; 374 AA.
XX
AC AAY25669;
XX
DT 30-SEP-1999 (first entry)
XX
DE Japanese cedar allergen 541802 Cry j I precursor protein fragment.
XX
KW Major histocompatibility complex; class II; desensitising; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX
OS Cedrus sp.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX
DR WPI; 1999-458255/38.
XX
PT Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 75; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitising patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chiromidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents a japanese cedar (Cedrus sp.) allergen 541802 Cry j I
CC precursor.
XX
SQ Sequence 374 AA;

Query Match 26.4%; Score 188; DB 20; Length 374;
Best Local Similarity 71.2%; Pred. No. 3.5e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 73 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGII 124
||||||||||||||||||||||||||||||||| |:| : ||:|| :|
Db 87 RPLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQVYIGNGGPCVFIKRVSNVI 138

RESULT 15
AAB69103
ID AAB69103 standard; Protein; 210 AA.
XX
AC AAB69103;
XX
DT 23-APR-2001 (first entry)
XX
DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:12.
XX
KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
KW T cell epitope; antisugipollinosis.
XX
OS Cryptomeria japonica.
OS Synthetic.
XX
PN JP2000327699-A.
XX
PD 28-NOV-2000.
XX
PF 15-MAR-2000; 2000JP-0071710.
XX
PR 15-MAR-1999; 99JP-0068316.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PA (SANY) SANKYO CO LTD.
XX
DR WPI; 2001-185061/19.
DR N-PSDB; AAF59012.

```
XX
PT   Novel peptide and its use -
XX
PS   Claim 7; Page 39-40; 75pp; Japanese.
XX
CC   The present invention describes a peptide, its complex, derivative or
CC   its polymerizate, where the peptide (I) has a formula of:
CC   alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
CC   where alpha1 to alpha7 = amino acid sequence selected from amino acid
CC   sequences ranging from 11-19 amino acids derived from T cell epitopes
CC   derived from cedar (japanese cedar - Cryptomeria japonica) pollen
CC   allergens. The peptide can be used in an antisugipollinosis agent.
CC   AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
CC   in the exemplification of the present invention.
XX
SQ   Sequence   210 AA;

  Query Match             25.0%;  Score 178;  DB 22;  Length 210;
  Best Local Similarity   32.5%;  Pred. No. 3e-14;
  Matches   53;  Conservative    7;  Mismatches    7;  Indels   96;  Gaps    6;

Qy       1 MKVTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNT-IGTGR--------- 50
           |||||||||||                     |||||||||||| : :|:
Db      29 MKVTVAFNQFGP--------------------FASKNFHLQKNTKLTSGKIASCLNYGL 67

Qy      51 ------------RWKNNRIWL--------QFAKLTGFTLMGRRPLWIIFSGNMNIKLKMP 90
                      :::   |:          ||||||||||||
Db      68 VHVANNNYDPSGKYEGGNIYTKKEAFNVEQFAKLTGFTLMGRA----------------- 110

Qy      91 MYIAGYKTFDGRRAEVSYVHVNGAKFIRRVDGIIAAYQNPASW 133
                 | |                   ||||||||||||
Db     111 ---------DPR------------------GIIAAYQNPASW 125


Search completed: May  5, 2003, 16:57:06
Job time : 47 secs
```

; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

```
Qy      45 TIGTGRRWKNNRIWLQFAKL------------TGF--TLMGRRPLWIIFSGNMN---IK 86
           : |||    |  : | |  |             :||   | | :|      |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
           | |  :  | |  |::  |    :    |  | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140
```

RESULT 15
US-10-174-590-58
; Sequence 58, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 58
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-58

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

```
Qy      45 TIGTGRRWKNNRIWLQFAKL------------TGF--TLMGRRPLWIIFSGNMN---IK 86
           : |||    |  : | |  |             :||   | | :|      |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
           | |  :  | |  |::  |    :    |  | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140
```

Search completed: May 5, 2003, 17:04:18
Job time : 49 secs

```
Qy      45 TIGTGRRWKNNRIWLQFAKL-------------TGF--TLMGRRPLWIIFSGNMN---IK 86
           : |||    |  : | |  |              :||   | | :|       |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
           | |  :  | |  |::  |    :    | | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140


RESULT 14
US-09-978-189-254
; Sequence 254, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C7
;  CURRENT APPLICATION NUMBER: US/09/978,189
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
```

; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

```
Qy      45 TIGTGRRWKNNRIWLQFAKL------------TGP--TLMGRRPLWIIFSGNMN---IK 86
           : |||    |  : | |  |                :||   | | :|      |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
             | |  :  | |  |:: |    :    |  | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140
```

RESULT 13
US-09-999-832A-254
; Sequence 254, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

```
Qy      45 TIGTGRRWKNNRIWLQFAKL-------------TGF--TLMGRRPLWIIFSGNMN---IK 86
           : |||    |  : | |  |                :||   | | :|      |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
             | |  :  | | |::  |    :    | | | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140
```

RESULT 12
US-09-978-192A-254
; Sequence 254, Application US/09978192A
; Patent No. US20020177553A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTIGN: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804

; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21

; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704

; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.7%; Score 62; DB 9; Length 545;
Best Local Similarity 25.7%; Pred. No. 47;
Matches 29; Conservative 11; Mismatches 49; Indels 24; Gaps 6;

```
Qy      45 TIGTGRRWKNNRIWLQFAKL-------------TGF--TLMGRRPLWIIFSGNMN---IK 86
           : |||     |  : | |  |               :||   | | :|      |  |   :
Db      29 SFGTGTS-SNPSVGLNFGNLGSTSTPATTSAPSSGFGTGLFGSKPATGFTLGGTNTGALH 87

Qy      87 LKMPMYIAGYKTFDGRRAEVSYVHVN---GAKFIRRVDGII--AAYQNPASWK 134
             | |  :  | |  |::  |    :    | | |   ||:  |  : |  ||
Db      88 TKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPPLGILRFAPPEPPEPWK 140
```

RESULT 11
US-09-978-697-254
; Sequence 254, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649

```
;   LENGTH: 1375
;   TYPE: PRT
;   ORGANISM: streptococcus mutans
US-09-740-274-4

  Query Match             8.8%;  Score 62.5;  DB 10;  Length 1375;
  Best Local Similarity   24.8%;  Pred. No. 1.3e+02;
  Matches   34;  Conservative   17;  Mismatches   53;  Indels   33;  Gaps    8;

Qy      10 FGPNRRVF---IKRVSNVIIHGRRIDIFASKNFHLQKNTIGTG-------RRWKNNRIWL 59
           ||   | |    |  |    :||  :  | :  |  :   | |          |   |:|
Db    1185 FGQQWRYFQNGIMAVGLTRVHG-AVQYFDASGFQAKGQFITTADGKLRYFDRDSGNQISN 1243

Qy      60 QFAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHVNG----AK 115
           :| :         :  | :|  |        : : |  ||:|:|    |   ||     :
Db    1244 RFVR-------NSKGEWFLFDHN-------GVAVTGTVTFNGQRL---YFKPNGVQAKGE 1286

Qy     116 FIRRVDGIIAAYQNPAS 132
           |||  :| :  | :| |
Db    1287 FIRDANGYL-RYYDPNS 1302


RESULT 10
US-09-978-295A-254
; Sequence 254, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C11
;  CURRENT APPLICATION NUMBER: US/09/978,295A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
```

```
US-10-121-032-15
; Sequence 15, Application US/10121032
; Patent No. US20020155550A1
;   GENERAL INFORMATION:
;        APPLICANT: Bylina, Edward J.
;        TITLE OF INVENTION: GLYCOSIDASE ENZYMES
;        NUMBER OF SEQUENCES: 72
;        CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Gray Cary Ware & Freidenrich LLP
;            STREET: 4365 Executive Drive, Suite 1600
;            CITY: San Diego
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 92121
;        COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: Windows95
;            SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/121,032
;            FILING DATE: 09-Apr-2002
;            CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/134,078
;            FILING DATE: 13-AUG-1998
;            APPLICATION NUMBER: 08/949,026
;            FILING DATE: 10-OCT-1997
;            APPLICATION NUMBER: 60/056,916
;            FILING DATE: 06-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;            NAME: Haile, Lisa A.
;            REGISTRATION NUMBER: 38,347
;            REFERENCE/DOCKET NUMBER: 09010/024002
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 858/677-1456
;            TELEFAX: 858/677-1465
;   INFORMATION FOR SEQ ID NO: 15:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 481 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-121-032-15

  Query Match             8.8%;  Score 62.5;  DB 9;  Length 481;
  Best Local Similarity   21.5%;  Pred. No. 35;
  Matches   41;  Conservative   28;  Mismatches   59;  Indels   63;  Gaps   11;

Qy       1 MKVTVAFNQFGP----NRRVFIKRVSN-VIIH-------GRRIDIFASKN-----FHLQK 43
           ::| | :::  |     | :| ::|  |  |:|          |:|  |:|         : |
Db      77 IRVGVEWSRIFPKPTFNVKVPVERDENGSIVHVDVDDKAVERLDELANKEAVNHYVEMYK 136

Qy      44 NTIGTGRR------------WKNNRI-------------WL------QFAKLTGFTL--M 70
           : :  ||:             | :| |              ||      :|||   :    |
Db     137 DWVERGRKLILNLYHWPLPLWLHNPIMVRRMGPDRAPSGWLNEESVVEFAKYAAYIAWKM 196

Qy      71 GRRP-LW--------IIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV----NGAKFI 117
           |  | :|         :   | | :|   |       :  |  |  :   |      |  :|
Db     197 GELPVMWSTMNEPNVVYEQGYMFVKGGFPPGYLSLEAADKARRNMIQAHARAYDNIKRFS 256

Qy     118 RRVDGIIAAYQ 128
           ::  |:| |:|
Db     257 KKPVGLIYAFQ 267


RESULT 8
US-10-093-037-15
; Sequence 15, Application US/10093037
; Publication No. US20030078397A1
; GENERAL INFORMATION:
;  APPLICANT: Jay M. Short
;  APPLICANT:  Bylina, Edward
;  APPLICANT:  Swanson, Ronald V.
;  APPLICANT:  Mathur, Eric J.
;  APPLICANT:  Lam, David E.
;  TITLE OF INVENTION: ENZYMES HAVING GLYCOSIDASE ACTIVITY AND METHODS OF USE THEREOF
;  FILE REFERENCE: 09010-024006
;  CURRENT APPLICATION NUMBER: US/10/093,037
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: US 09/910,579
;  PRIOR FILING DATE: 2001-07-20
;  PRIOR APPLICATION NUMBER: US 09/134,078
;  PRIOR FILING DATE: 1998-08-13
;  PRIOR APPLICATION NUMBER: US 08/949,026
;  PRIOR FILING DATE: 1997-10-10
;  PRIOR APPLICATION NUMBER: US 60/056,916
;  PRIOR FILING DATE: 1996-12-06
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Desulfurococcus sp.
US-10-093-037-15

  Query Match             8.8%;  Score 62.5;  DB 9;  Length 481;
  Best Local Similarity   21.5%;  Pred. No. 35;
  Matches   41;  Conservative   28;  Mismatches   59;  Indels   63;  Gaps   11;

Qy       1 MKVTVAFNQFGP----NRRVFIKRVSN-VIIH-------GRRIDIFASKN-----FHLQK 43
           ::| | :::  |     | :| ::|  |  |:|          |:|  |:|         : |
Db      77 IRVGVEWSRIFPKPTFNVKVPVERDENGSIVHVDVDDKAVERLDELANKEAVNHYVEMYK 136

Qy      44 NTIGTGRR------------WKNNRI-------------WL------QFAKLTGFTL--M 70
           : :  ||:             | :| |              ||      :|||   :    |
Db     137 DWVERGRKLILNLYHWPLPLWLHNPIMVRRMGPDRAPSGWLNEESVVEFAKYAAYIAWKM 196

Qy      71 GRRP-LW--------IIFSGNMNIKLKMPMYIAGYKTFDGRRAEVSYVHV----NGAKFI 117
           |  | :|         :   | | :|   |       :  |  |  :   |      |  :|
Db     197 GELPVMWSTMNEPNVVYEQGYMFVKGGFPPGYLSLEAADKARRNMIQAHARAYDNIKRFS 256

Qy     118 RRVDGIIAAYQ 128
           ::  |:| |:|
Db     257 KKPVGLIYAFQ 267


RESULT 9
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
;  APPLICANT:  Nichols, Scott E.
;  TITLE OF INVENTION:  Glucan-containing Compositions and Paper
;  FILE REFERENCE:  0357CRD
;  CURRENT APPLICATION NUMBER: US/09/740,274
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER:  09/210,361
;  PRIOR FILING DATE:  1998-12-11
;  PRIOR APPLICATION NUMBER:  09/007,999
;  PRIOR FILING DATE:  1998-01-16
;  PRIOR APPLICATION NUMBER:  08/478,704
;  PRIOR FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER:  09/009,620
;  PRIOR FILING DATE:  1998-01-20
;  PRIOR APPLICATION NUMBER:  08/485,243
;  PRIOR FILING DATE:  1995-06-07
;  PRIOR APPLICATION NUMBER:  09/008,172
;  PRIOR FILING DATE:  1998-01-16
;  PRIOR APPLICATION NUMBER:  08/482,711
;  PRIOR FILING DATE:  1995-06-07
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  Adler, Jon Elliot
;  APPLICANT:  Ryba, Nick
;  APPLICANT:  Mueller, Ken
;  APPLICANT:  Hoon, Mark
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION:  T2R, a No. US20030022278A1el Family of Taste Receptors
;  FILE REFERENCE:  02307E-098010US
;  CURRENT APPLICATION NUMBER: US/09/510,332
;  CURRENT FILING DATE:  2000-02-22
;  PRIOR APPLICATION NUMBER:  US 09/393,634
;  PRIOR FILING DATE:  1999-09-10
;  NUMBER OF SEQ ID NOS:  172
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 85
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Rattus sp.
;   FEATURE:
;   OTHER INFORMATION: rat T2R05 (rGR05)
US-09-510-332-85

  Query Match             8.9%;  Score 63;  DB 9;  Length 309;
  Best Local Similarity   28.4%;  Pred. No. 18;
  Matches   23;  Conservative   15;  Mismatches   33;  Indels   10;  Gaps    4;

Qy      52 WKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKL---KMPMYIAGYKTFDGRRAE--V 106
           |   || : || | |  ||     | :||  : :|:   |  :||        :::|  :
Db     121 WLKRRINIVFAFLIGCLLMS----W-LFSFPVVVKMVKDKKMLYINSSWQIHMKKSELII 175

Qy     107 SYVHVNGAKFIRRVDGIIAAY 127
           :||  ||  |:  :  :|  :
Db     176 NYVFTNGGVFLLFIIMLIVCF 196


RESULT 5
US-08-915-048A-2
; Sequence 2, Application US/08915048A
; Patent No. US20020025569A1
;  GENERAL INFORMATION:
;    APPLICANT:  Caligiuri, Maureen
;    APPLICANT:  Rolfe, Mark
;    TITLE OF INVENTION:  Components of Ubiquitin Ligase Complexes
;    TITLE OF INVENTION:  and Uses Related Thereto
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley, Hoag & Eliot
;      STREET:  One Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII (text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,048A
;      FILING DATE:  20-AUG-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MIV-092.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 832-1000
;      TELEFAX:  (617) 832-7000
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1121 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-915-048A-2

  Query Match             8.9%;  Score 63;  DB 8;  Length 1121;
  Best Local Similarity   25.2%;  Pred. No. 86;
  Matches   26;  Conservative   14;  Mismatches   23;  Indels   40;  Gaps    6;

Qy      10 FGPNRRVFIKRVSNV----IIHGRR-------------IDIFASKNFHLQKNTIGTGRR 51
           || : |: |||| |:    :||| |             : :| ||   : | : | |
Db      46 FGGHLRM-IKRVQNLLGHYLIHGFRVRPEPNGDLDLEAMVAVFGSKGLRVVKISWGQGHF 104

Qy      52 WK-------NNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKL 87
           |:       |   |:  |:           |:   ||: : |
Db     105 WELWRSGLWNMSDWIWDAR----------WL--EGNIALAL 133


RESULT 6
US-10-114-893-73
; Sequence 73, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT: McCoy, John M.
;  APPLICANT: LaVallie, Edward R.
;  APPLICANT: Collins-Racie, Lisa A.
;  APPLICANT: Evans, Cheryl
;  APPLICANT: Merberg, David
;  APPLICANT: Treacy, Maurice
;  APPLICANT: Bowman, Michael R.
;  APPLICANT: Spaulding, Vikki
;  APPLICANT: Carlin-Duckett, McKeough
;  APPLICANT: Kelleher, Kerry S.
;  APPLICANT: Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: GI 6000-10A
;  CURRENT APPLICATION NUMBER: US/10/114,893
;  CURRENT FILING DATE: 2002-04-02
;  EARLIER APPLICATION NUMBER: 09/413,232
;  EARLIER FILING DATE: 1999-10-06
;  NUMBER OF SEQ ID NOS: 321
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
;   LENGTH: 380
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (135)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (179)
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (318)
US-10-114-893-73

  Query Match             8.8%;  Score 62.5;  DB 9;  Length 380;
  Best Local Similarity   27.4%;  Pred. No. 26;
  Matches   20;  Conservative   14;  Mismatches   32;  Indels    7;  Gaps    3;

Qy      52 WKN--NRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKLKMPMYIA--GYKTFDGRRAEVS 107
           |||  | |:: || :   | :   | ||:   :  :   : :| |  ||  |:     :
Db     257 WKNTCNNIFIVFAIVFIITRLVILPFWIL---HCTLVYPLELYPAFFGYYFFNSMMGVLQ 313

Qy     108 YVHVNGAKFIRRV 120
           :|:  |  | |:
Db     314 LLHIXWAYLILRM 326


RESULT 7
```

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (415) 705-8410
;              TELEFAX: (415) 397-8338
;    INFORMATION FOR SEQ ID NO: 7:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 811 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: unknown
;              TOPOLOGY: unknown
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-011-366-7

  Query Match             10.0%;  Score 71;  DB 9;  Length 811;
  Best Local Similarity   21.7%;  Pred. No. 6.6;
  Matches   36;  Conservative   22;  Mismatches   64;  Indels   44;  Gaps    7;

Qy       3 VTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQFA 62
           :|:  | |  :     | |: |       : || | |   | |   ::|  : |
Db     383 ITIERNNFYFDANNESKMVTGVFKGPNGFEYFAPANTH-NNNIEGQAIVYQNKFLTLNGK 441

Qy      63 K---------LTGF-TLMGRRPLWIIFSGNMN----------------IKLKMPMYIAG 95
           |         :||: |: |::     :  |:|               |       |
Db     442 KYYFDNDSKAVTGWQTIDGKK-----YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATG 496

Qy      96 YKTFDGRR---------AEVSYVHVNGAKFIRRVDGI--IAAYQNP 130
           ::| ||::         |   |  :||  |    |||  |  :: |
Db     497 WQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGP 542


RESULT 2
US-10-011-366-6
; Sequence 6, Application US/10011366
; Publication No. US20030054493A1
;   GENERAL INFORMATION:
;         APPLICANT: Williams, James A.
;                    Kink, John A.
;         TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
;                             OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
;                             DISEASE
;         NUMBER OF SEQUENCES: 22
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Medlen & Carroll
;              STREET: 220 Montgomery Street, Suite 2200
;              CITY: San Francisco
;              STATE: California
;              COUNTRY: United States of America
;              ZIP: 94104
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/011,366
;              FILING DATE: 16-No. US20030054493A1-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US/08/957,310
;              FILING DATE: 23-OCT-1997
;              APPLICATION NUMBER: US 08/329,154
;              FILING DATE: 24-OCT-1994
;              APPLICATION NUMBER: US 08/161,907
;              FILING DATE: 02-DEC-1993
;              APPLICATION NUMBER: US 07/985,321
;              FILING DATE: 04-DEC-1992
;              APPLICATION NUMBER: US 07/429,791
;              FILING DATE: 31-OCT-1989
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Ingolia, Diane E.
;              REGISTRATION NUMBER: 40,027
;              REFERENCE/DOCKET NUMBER: OPHD-01121
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (415) 705-8410
;              TELEFAX: (415) 397-8338
;    INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 2710 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-011-366-6

  Query Match             10.0%;  Score 71;  DB 9;  Length 2710;
  Best Local Similarity   21.7%;  Pred. No. 29;
  Matches   36;  Conservative   22;  Mismatches   64;  Indels   44;  Gaps    7;

Qy       3 VTVAFNQFGPNRRVFIKRVSNVIIHGRRIDIFASKNFHLQKNTIGTGRRWKNNRIWLQFA 62
           :|:  | |  :     | |: |       : || | |   | |   ::|  : |
Db    2256 ITIERNNFYFDANNESKMVTGVFKGPNGFEYFAPANTH-NNNIEGQAIVYQNKFLTLNGK 2314

Qy      63 K---------LTGF-TLMGRRPLWIIFSGNMN----------------IKLKMPMYIAG 95
           |         :||: |: |::     :  |:|               |       |
Db    2315 KYYFDNDSKAVTGWQTIDGKK-----YYFNLNTAEAATGWQTIDGKKYYFNLNTAEAATG 2369

Qy      96 YKTFDGRR---------AEVSYVHVNGAKFIRRVDGI--IAAYQNP 130
           ::| ||::         |   |  :||  |    |||  |  :: |
Db    2370 WQTIDGKKYYFNTNTFIASTGYTSINGKHFYFNTDGIMQIGVFKGP 2415


RESULT 3
US-09-393-634-9
; Sequence 9, Application US/09393634
; Patent No. US20020051997A1
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  Adler, Jon Elliot
;  APPLICANT:  Ryba, Nick
;  APPLICANT:  Mueller, Ken
;  APPLICANT:  Hoon, Mark
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION:  SF, a No. US20020051997A1el Family of Taste Receptors
;  FILE REFERENCE:  02307E-098000US
;  CURRENT APPLICATION NUMBER: US/09/393,634
;  CURRENT FILING DATE:  1999-09-10
;  NUMBER OF SEQ ID NOS:  92
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 224
;   TYPE: PRT
;   ORGANISM: Rattus sp.
;   FEATURE:
;   OTHER INFORMATION: rat GR05
US-09-393-634-9

  Query Match             9.0%;  Score 64;  DB 10;  Length 224;
  Best Local Similarity   28.4%;  Pred. No. 9.2;
  Matches   23;  Conservative   15;  Mismatches   33;  Indels   10;  Gaps    4;

Qy      52 WKNNRIWLQFAKLTGFTLMGRRPLWIIFSGNMNIKL---KMPMYIAGYKTFDGRRAE--V 106
           |   || : || | |  ||     | :||  : :|:   |  :||        :::|  :
Db     121 WLKRRINIVFAFLIGCLLMS----W-LFSFPVVVKMVKDKKMLYINSSWQIHMKKSELII 175

Qy     107 SYVHVNGAKFIRRVDGIIAAY 127
           :||  ||  |:  :  :|  :
Db     176 NYVFTNGGVFLLFIIMVIGCF 196


RESULT 4
US-09-510-332-85
; Sequence 85, Application US/09510332
; Publication No. US20030022278A1
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 5, 2003, 16:59:07 ; Search time 46 Seconds
(without alignments)
251.357 Million cell updates/sec

Title: US-09-142-524D-3
Perfect score: 711
Sequence: 1 MKVTVAFNQFGPNRRVFIKR.....KFIRRVDGIIAAYQNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	10.0	811	9	US-10-011-366-7	Sequence 7, Appli
2	71	10.0	2710	9	US-10-011-366-6	Sequence 6, Appli
3	64	9.0	224	10	US-09-393-634-9	Sequence 9, Appli
4	63	8.9	309	9	US-09-510-332-85	Sequence 85, Appl
5	63	8.9	1121	8	US-08-915-048A-2	Sequence 2, Appli
6	62.5	8.8	380	9	US-10-114-893-73	Sequence 73, Appl
7	62.5	8.8	481	9	US-10-121-032-15	Sequence 15, Appl
8	62.5	8.8	481	9	US-10-093-037-15	Sequence 15, Appl
9	62.5	8.8	1375	10	US-09-740-274-4	Sequence 4, Appli
10	62	8.7	545	9	US-09-978-295A-254	Sequence 254, App
11	62	8.7	545	9	US-09-978-697-254	Sequence 254, App
12	62	8.7	545	9	US-09-978-192A-254	Sequence 254, App
13	62	8.7	545	9	US-09-999-832A-254	Sequence 254, App
14	62	8.7	545	9	US-09-978-189-254	Sequence 254, App
15	62	8.7	545	9	US-10-174-590-58	Sequence 58, Appl
16	62	8.7	545	9	US-10-176-758-58	Sequence 58, Appl
17	62	8.7	545	9	US-10-175-737-58	Sequence 58, Appl
18	62	8.7	545	9	US-10-173-706-58	Sequence 58, Appl
19	62	8.7	545	9	US-10-175-738-58	Sequence 58, Appl
20	62	8.7	545	9	US-10-175-752-58	Sequence 58, Appl
21	62	8.7	545	9	US-10-176-482-58	Sequence 58, Appl
22	62	8.7	545	9	US-10-176-757-58	Sequence 58, Appl
23	62	8.7	545	9	US-10-176-913-58	Sequence 58, Appl
24	62	8.7	545	9	US-10-180-552-58	Sequence 58, Appl
25	62	8.7	545	9	US-10-180-557-58	Sequence 58, Appl
26	62	8.7	545	9	US-10-173-700-58	Sequence 58, Appl
27	62	8.7	545	9	US-10-174-572-58	Sequence 58, Appl
28	62	8.7	545	9	US-10-174-579-58	Sequence 58, Appl
29	62	8.7	545	9	US-10-174-582-58	Sequence 58, Appl
30	62	8.7	545	9	US-10-174-588-58	Sequence 58, Appl
31	62	8.7	545	9	US-10-175-739-58	Sequence 58, Appl
32	62	8.7	545	9	US-10-175-740-58	Sequence 58, Appl
33	62	8.7	545	9	US-10-175-743-58	Sequence 58, Appl
34	62	8.7	545	9	US-10-176-488-58	Sequence 58, Appl
35	62	8.7	545	9	US-10-176-492-58	Sequence 58, Appl
36	62	8.7	545	9	US-10-176-747-58	Sequence 58, Appl
37	62	8.7	545	9	US-10-176-750-58	Sequence 58, Appl
38	62	8.7	545	9	US-10-176-985-58	Sequence 58, Appl
39	62	8.7	545	9	US-10-176-987-58	Sequence 58, Appl
40	62	8.7	545	9	US-10-176-991-58	Sequence 58, Appl
41	62	8.7	545	9	US-10-176-992-58	Sequence 58, Appl
42	62	8.7	545	9	US-10-176-993-58	Sequence 58, Appl
43	62	8.7	545	9	US-10-184-658-58	Sequence 58, Appl
44	62	8.7	545	9	US-10-173-695-58	Sequence 58, Appl
45	62	8.7	545	9	US-10-173-697-58	Sequence 58, Appl

ALIGNMENTS

```
RESULT 1
US-10-011-366-7
; Sequence 7, Application US/10011366
; Publication No. US20030054493A1
;  GENERAL INFORMATION:
;       APPLICANT: Williams, James A.
;                  Kink, John A.
;       TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
;                           OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
;                           DISEASE
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/011,366
;            FILING DATE: 16-No. US20030054493A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/957,310
;            FILING DATE: 23-OCT-1997
;            APPLICATION NUMBER: US 08/329,154
;            FILING DATE: 24-OCT-1994
;            APPLICATION NUMBER: US 08/161,907
;            FILING DATE: 02-DEC-1993
;            APPLICATION NUMBER: US 07/985,321
;            FILING DATE: 04-DEC-1992
;            APPLICATION NUMBER: US 07/429,791
;            FILING DATE: 31-OCT-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: OPHD-01121
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 12.6711 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-152
Perfect score: 67
Sequence: 1 LSDISLKLTSGKIAS 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	51	76.1	507	10	Q9FY19	Q9fy19 juniperus a
2	45	67.2	251	16	Q8Z5W6	Q8z5w6 salmonella
3	45	67.2	268	16	Q8ZNV7	Q8znv7 salmonella
4	44	65.7	242	16	Q97N70	Q97n70 streptococc
5	43	64.2	275	16	Q9KKE1	Q9kke1 rhizobium m
6	43	64.2	717	11	Q9CXJ4	Q9cxj4 mus musculu
7	42	62.7	261	2	Q9R6G1	Q9r6g1 agrobacteri
8	42	62.7	1376	3	Q8X1P2	Q8x1p2 podospora a
9	40	59.7	134	16	Q8UDM0	Q8udm0 agrobacteri
10	40	59.7	253	16	Q8ZEU3	Q8zeu3 yersinia pe
11	40	59.7	264	16	Q8RGI6	Q8rgi6 fusobacteri
12	40	59.7	278	9	Q8SCM8	Q8scm8 pseudomonas
13	40	59.7	299	16	Q99RT2	Q99rt2 staphylococ
14	40	59.7	299	16	Q931F9	Q931f9 staphylococ
15	40	59.7	317	3	Q9UUD9	Q9uud9 schizosacch
16	40	59.7	1066	4	Q96DN5	Q96dn5 homo sapien
17	39	58.2	93	16	Q9K9U3	Q9k9u3 bacillus ha
18	39	58.2	225	2	Q9AHL4	Q9ahl4 pasteurella
19	39	58.2	292	16	Q8ZNG8	Q8zng8 salmonella
20	39	58.2	292	16	Q8Z557	Q8z557 salmonella
21	39	58.2	325	10	Q9SZI9	Q9szi9 arabidopsis
22	39	58.2	393	16	O85461	O85461 pasteurella
23	39	58.2	649	10	Q8W1C1	Q8w1c1 brassica ju
24	39	58.2	839	10	Q9M0G3	Q9m0g3 arabidopsis
25	39	58.2	1446	10	Q9SVT9	Q9svt9 arabidopsis
26	38	56.7	226	16	Q92AT4	Q92at4 listeria in
27	38	56.7	226	16	Q8Y6G3	Q8y6g3 listeria mo
28	38	56.7	230	16	Q9RC35	Q9rc35 bacillus ha
29	38	56.7	241	16	O30650	O30650 vibrio chol
30	38	56.7	251	16	Q9PMU5	Q9pmu5 campylobact
31	38	56.7	257	16	Q8XWN5	Q8xwn5 ralstonia s
32	38	56.7	273	16	Q9KU04	Q9ku04 vibrio chol
33	38	56.7	291	16	Q8ZKC9	Q8zkc9 salmonella
34	38	56.7	291	16	Q8Z1M7	Q8z1m7 salmonella
35	38	56.7	359	13	Q9PVY7	Q9pvy7 anguilla an
36	38	56.7	398	10	Q9SJR5	Q9sjr5 arabidopsis
37	38	56.7	475	16	Q98GH7	Q98gh7 rhizobium l
38	38	56.7	505	16	Q98QA7	Q98qa7 mycoplasma
39	38	56.7	529	12	O39281	O39281 equine herp
40	38	56.7	548	10	Q9LFA0	Q9lfa0 arabidopsis
41	38	56.7	568	16	Q8XK20	Q8xk20 clostridium
42	37.5	56.0	1167	12	Q8QV02	Q8qv02 fiji diseas
43	37	55.2	124	16	O34380	O34380 bacillus su
44	37	55.2	165	16	Q99VW8	Q99vw8 staphylococ
45	37	55.2	194	2	Q9KI70	Q9ki70 bifidobacte

ALIGNMENTS

RESULT 1
Q9FY19
ID Q9FY19 PRELIMINARY; PRT; 507 AA.
AC Q9FY19;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pollen major allergen 2 protein precursor.
GN JNA2.
OS Juniperus ashei (Ozark white cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxID=13101;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MALE POLLEN;
RX MEDLINE=20403896; PubMed=10944464;
RA Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;
RT "Purification, Identification and cDNA cloning of Jun a 2, the second
RT major allergen of mountain cedar pollen.";
RL Biochem. Biophys. Res. Commun. 275:195-202(2000).
CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC (POLYGALACTURONASES).
DR EMBL; AJ404653; CAC05582.1; -.
DR HSSP; P26509; 1BHE.
DR InterPro; IPR000743; GH28.
DR InterPro; IPR000408; Reg_chr_condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW Cell wall; Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 54 POTENTIAL.
SQ SEQUENCE 507 AA; 55730 MW; 2B2E0AA5E958FE5A CRC64;

Query Match 76.1%; Score 51; DB 10; Length 507;
Best Local Similarity 66.7%; Pred. No. 0.58;
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSDISLKLTSGKIAS 15

```
          ||:: |||||||:|:
Db    396 LSNVFLKLTSGKVAT 410


RESULT 2
Q8Z5W6
ID   Q8Z5W6      PRELIMINARY;      PRT;   251 AA.
AC   Q8Z5W6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   High-affinity zinc uptake system ATP-binding protein.
GN   STY2100.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL627272; CAD05643.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   251 AA;  27731 MW;  4099397813CBBBE0 CRC64;

  Query Match             67.2%;  Score 45;  DB 16;  Length 251;
  Best Local Similarity   69.2%;  Pred. No. 3.7;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |||:||:|: |||
Db      20 LSDVSLELSPGKI 32


RESULT 3
Q8ZNV7
ID   Q8ZNV7      PRELIMINARY;      PRT;   268 AA.
AC   Q8ZNV7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   ABC superfamily (atp_bind) high affinity Zn transport protein.
GN   ZNUC OR STM1892.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008784; AAL20808.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
KW   Complete proteome.
SQ   SEQUENCE   268 AA;  29750 MW;  D5669D2448A74B28 CRC64;

  Query Match             67.2%;  Score 45;  DB 16;  Length 268;
  Best Local Similarity   69.2%;  Pred. No. 4;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |||:||:|: |||
Db      37 LSDVSLELSPGKI 49


RESULT 4
Q97N70
ID   Q97N70      PRELIMINARY;      PRT;   242 AA.
AC   Q97N70;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   ABC transporter, ATP-binding protein.
GN   SP2196.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007508; AAK76247.1; -.
DR   TIGR; SP2196; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
DR   PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   242 AA;  27519 MW;  F88BD2FDD31DDB35 CRC64;

  Query Match             65.7%;  Score 44;  DB 16;  Length 242;
  Best Local Similarity   53.3%;  Pred. No. 5.5;
  Matches    8;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKIAS 15
           | ||:|::|||:: |
Db      19 LEDINLQVTSGEVVS 33


RESULT 5
Q9KKE1
ID   Q9KKE1      PRELIMINARY;      PRT;   275 AA.
AC   Q9KKE1;
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Inner membrane ATPase protein (Histidine transport ATP-binding ABC
DE   transporter protein).
GN   HISV OR R02702 OR SMC00670.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RM5000;
RX   MEDLINE=20309701; PubMed=10850986;
RA   Boncompagni E., Dupont L., Mignot T., Osteras M., Lambert A.,
RA   Poggi M.-C., Le Rudulier D.;
RT   "Characterization of a Sinorhizobium meliloti ATP-Binding Cassette
RT   Histidine Transporter Also Involved in Betaine and Proline Uptake.";
RL   J. Bacteriol. 182:3717-3725(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR   EMBL; AF111939; AAF78099.1; -.
DR   EMBL; AL591791; CAC47281.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport; Complete proteome.
SQ   SEQUENCE   275 AA;  30364 MW;  99C41F56CF93790F CRC64;

  Query Match             64.2%;  Score 43;  DB 16;  Length 275;
  Best Local Similarity   61.5%;  Pred. No. 9.6;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |:|:|||: :|||
Db      43 LNDVSLKIGAGKI 55


RESULT 6
Q9CXJ4
ID   Q9CXJ4      PRELIMINARY;      PRT;   717 AA.
AC   Q9CXJ4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   14, 17 days embryo head cDNA, RIKEN full-length enriched library,
DE   clone:3222401P09, full insert sequence (RIKEN cDNA 4833412N02
DE   gene).
GN   4833412N02RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK014319; BAB29270.1; -.
DR   EMBL; BC015301; AAH15301.1; -.
DR   HSSP; P13569; 1NBD.
DR   MGD; MGI:1921860; 4833412N02Rik.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABCtranprtrTM.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding.
SQ   SEQUENCE   717 AA;  77999 MW;  E573D4276B9658EB CRC64;

  Query Match             64.2%;  Score 43;  DB 11;  Length 717;
  Best Local Similarity   60.0%;  Pred. No. 25;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKIAS 15
           | | :||| |||| :
Db     472 LKDFTLKLPSGKIVA 486


RESULT 7
Q9R6G1
ID   Q9R6G1      PRELIMINARY;      PRT;   261 AA.
AC   Q9R6G1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   TIORF98 protein.
GN   TIORF98.
OS   Agrobacterium tumefaciens.
OG   Plasmid pTi-SAKURA.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=358;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RX   MEDLINE=20184752; PubMed=10721727;
RA   Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA   Katoh A., Yoshida K.;
RT   "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";
RL   Gene 242:331-336(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RX   MEDLINE=98193120; PubMed=9524202;
```

```
RA   Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT   "Novel structural difference between nopaline- and octopine- type trbJ
RT   gene:construction of genetic and physical map and sequencing of
RT   trb/traI and rep gene clusters of a new Ti plasmid pTi-SAKURA.";
RL   Biochim. Biophys. Acta 1396:1-7(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RA   Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT   "Genome structure of pTi-SAKURA(I): Strategy for DNA sequencing of a
RT   Japanese cherry-Ti plasmid.";
RL   Nucleic Acids Symp. Ser. 37:159-160(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RA   Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT   "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA.";
RL   Nucleic Acids Symp. Ser. 39:185-186(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RA   Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
RT   "Genome structure of pTi-SAKURA (IV): Characteristics of tra region.";
RL   Nucleic Acids Symp. Ser. 39:187-188(1998).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF301001;
RA   Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT   "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
RT   plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens.";
RL   Nucleic Acids Symp. Ser. 39:265-266(1998).
DR   EMBL; AB016260; BAA87723.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Plasmid.
SQ   SEQUENCE   261 AA;  27971 MW;  451054F02082BF9B CRC64;

  Query Match             62.7%;  Score 42;  DB 2;  Length 261;
  Best Local Similarity   61.5%;  Pred. No. 14;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |||:|||  :||:
Db      35 LSDVSLKAWAGKV 47


RESULT 8
Q8X1P2
ID   Q8X1P2      PRELIMINARY;      PRT;  1376 AA.
AC   Q8X1P2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta transducin-like protein HET-D2Y.
GN   HET-D.
OS   Podospora anserina.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Lasiosphaeriaceae; Podospora.
OX   NCBI_TaxID=5145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Espagne E., Balhadere P., Begueret J., Turcq B.;
RT   "A new family of WD40 proteins implicated in vegetative
RT   incompatibility; evidence for a major role of WD40 repeat domain in
RT   the specificity of het-e and het-d genes.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF323585; AAL37301.1; -.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 12.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   ProDom; PD000018; WD40; 9.
DR   SMART; SM00320; WD40; 12.
DR   PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_11.
DR   PROSITE; PS50082; WD_REPEATS_2; 11.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
SQ   SEQUENCE   1376 AA;  152068 MW;  2191826AA5C62D97 CRC64;

  Query Match             62.7%;  Score 42;  DB 3;  Length 1376;
  Best Local Similarity   69.2%;  Pred. No. 73;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           :|| |||| |||:
Db     643 ISDRSLKLLSGKL 655


RESULT 9
Q8UDM0
ID   Q8UDM0      PRELIMINARY;      PRT;  134 AA.
AC   Q8UDM0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein Atu2104.
GN   ATU2104 OR AGR_C_3816.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE009161; AAL43095.1; ALT_INIT.
DR   EMBL; AE008127; AAK87853.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   134 AA;  14071 MW;  EB868269D7DDF8F5 CRC64;

  Query Match             59.7%;  Score 40;  DB 16;  Length 134;
  Best Local Similarity   40.0%;  Pred. No. 17;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKIAS 15
           ::|: ||:| | :|:
Db      99 MTDLHLKITEGNVAA 113


RESULT 10
Q8ZEU3
```

```
ID   Q8ZEU3      PRELIMINARY;      PRT;   253 AA.
AC   Q8ZEU3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   High-affinity zinc uptake system ATP-binding protein.
GN   ZNUC OR YPO2060.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
DR   EMBL; AJ414151; CAC90872.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   253 AA;  27620 MW;  15D2023879279169B CRC64;

  Query Match             59.7%;  Score 40;  DB 16;  Length 253;
  Best Local Similarity   69.2%;  Pred. No. 32;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |:|||| |  |||
Db      21 LNDISLSLRPGKI 33


RESULT 11
Q8RGI6
ID   Q8RGI6      PRELIMINARY;      PRT;   264 AA.
AC   Q8RGI6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Iron(III) dicitrate transport ATP-binding protein fecE.
GN   FN0307.
OS   Fusobacterium nucleatum (subsp. nucleatum).
OC   Bacteria; Fusobacteria; Fusobacterium.
OX   NCBI_TaxID=76856;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25586;
RX   MEDLINE=21886394; PubMed=11889109;
RA   Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA   Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA   Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA   Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA   Fonstein M., Kyrpides N., Overbeek R.;
RT   "Genome sequence and analysis of the oral bacterium Fusobacterium
RT   nucleatum strain ATCC 25586.";
RL   J. Bacteriol. 184:2005-2018(2002).
DR   EMBL; AE010543; AAL94513.1; -.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   264 AA;  30009 MW;  04C8BDA94F121B53 CRC64;

  Query Match             59.7%;  Score 40;  DB 16;  Length 264;
  Best Local Similarity   61.5%;  Pred. No. 33;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           | ||:||: ||:|
Db      22 LKDINLKIDSGEI 34


RESULT 12
Q8SCM8
ID   Q8SCM8      PRELIMINARY;      PRT;   278 AA.
AC   Q8SCM8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   PHIKZ284.
OS   Pseudomonas phage phiKZ.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX   NCBI_TaxID=169683;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21914557; PubMed=11916376;
RA   Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA   Bourkaltseva M.V., Sykilinda N.N., Krylov V.N., Volckaert G.;
RT   "The genome of bacteriophage phiKZ of Pseudomonas aeruginosa.";
RL   J. Mol. Biol. 317:1-19(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA   Bourkaltseva M.V., Sykilinda N.N., Krylov V.V., Volckaert G.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF399011; AAL83185.1; -.
SQ   SEQUENCE   278 AA;  31883 MW;  C7A2CA9FDE57AD5A CRC64;

  Query Match             59.7%;  Score 40;  DB 9;  Length 278;
  Best Local Similarity   53.8%;  Pred. No. 35;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 DISLKLTSGKIAS 15
           |: :|||:||: |
Db     220 DVLIKLTNGKVYS 232


RESULT 13
Q99RT2
ID   Q99RT2      PRELIMINARY;      PRT;   299 AA.
AC   Q99RT2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein SA2132.
GN   SA2132.
OS   Staphylococcus aureus (strain N315).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Staphylococcus.
OX   NCBI_TaxID=158879;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA   Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA   Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA   Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA   Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA   Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003136; BAB43433.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   SMART; SM00382; AAA; 1.
```

```
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   299 AA;  33723 MW;  BA8FEC5B011FFD7C CRC64;

  Query Match             59.7%;  Score 40;  DB 16;  Length 299;
  Best Local Similarity   61.5%;  Pred. No. 37;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           ::||||:| |||:
Db      18 VNDISLELESGKM 30


RESULT 14
Q931F9
ID   Q931F9      PRELIMINARY;      PRT;   299 AA.
AC   Q931F9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein SAV2341.
GN   SAV2341.
OS   Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Staphylococcus.
OX   NCBI_TaxID=158878;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003365; BAB58503.1; -.
DR   InterPro; IPR003439; ABC_transportr.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transportr; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   299 AA;  33757 MW;  798FEC58F21FFD7A CRC64;

  Query Match             59.7%;  Score 40;  DB 16;  Length 299;
  Best Local Similarity   61.5%;  Pred. No. 37;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           ::||||:| |||:
Db      18 VNDISLELESGKM 30


RESULT 15
Q9UUD9
ID   Q9UUD9      PRELIMINARY;      PRT;   317 AA.
AC   Q9UUD9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Hypothetical 37.0 kDa protein.
GN   SPBC17G9.13C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   Wood V., Rajandream M.A., Barrell B.G., Volckaert G.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL109846; CAB52811.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   317 AA;  37009 MW;  C4356E2E97057126 CRC64;

  Query Match             59.7%;  Score 40;  DB 3;  Length 317;
  Best Local Similarity   61.5%;  Pred. No. 40;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LSDISLKLTSGKI 13
           |||| |:|  ||:
Db     297 LSDIQLRLNVGKV 309


Search completed: April 20, 2003, 13:13:23
Job time : 14.6711 secs
```